

```
RESULT 2
S43467
nef protein - human immunodeficiency virus type 1
C:Species: human immunodeficiency virus type 1, HIV-1
C:Date: 20-Oct-1994 #sequence_revision 10-Nov-1995 #text_change 03-May-1996
C:Accession: S43467
R:Freund, J.; Kellner, R.; Houthaeve, T.; Kalbitzer, H.R.
Eur. J. Biochem. 221, 811-819, 1994
A:Title: Stability and proteolytic domains of Nef protein from human immunodeficiency vi
A:Reference number: S43467; MUID:94229079
A:Accession: S43467
A>Status: preliminary
A:Molecule type: protein
A:Residues: 1-205 <FRET>
C:Superfamily: AIDS nef protein

Query Match 66.1%; Score 1115; DB 2; Length 205;
Best Local Similarity 99.0%; Pred. No. 3.2e-85;
Matches 203; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 2 GKGKSSVVGWPTVRRMRRAEPAADGGAASRDLEKHGAISSNTAATNAACAWLEA 61
|||||
DB 1 GKGKSSVVGWPAVRMRRAEPAADGGAASRDLEKHGAISSNTAATNAACAWLEA 60
|||||
QY 62 EEEVGFPVTPQVPLRPMTYKAADVLSHFLKEKGLGLIHSQRQDILDLYHTQGYF 121
|||||
DB 61 EEEVGFPVTPQVPLRPMTYKAADVLSHFLKEKGLGLIHSQRQDILDLYHTQGYF 120
|||||
QY 122 PDQWNTYTPGQVRYPLTFGWCYKLVPEPDKVEEANKGENTSLHPVSLHGMDDPEREVL 181
|||||
DB 121 PDQWNTYTPGQVRYPLTFGWCYKLVPEPDKVEEANKGENTSLHPVSLHGMDDPEREVL 180
|||||
QY 182 EWRFDLSRLAFHHVARELHPEYFKNC 206
|||||
DB 181 EWRFDLSRLAFHHVARELHPEYFKNC 205
|||||

RESULT 3
S03244
nef protein (clone HX82) - human immunodeficiency virus type 1
N:Alternate names: 3'-orf protein
C:Species: human immunodeficiency virus type 1, HIV-1
C:Date: 28-Feb-1990 #sequence_revision 28-Feb-1990 #text_change 30-Sep-1993
C:Accession: S03244
R:Ratner, L.; Starcich, B.; Josephs, S.F.; Hahn, B.H.; Reddy, E.P.; Livak, K.J.; Pettawa
Nucleic Acids Res. 13, 8219-8229, 1985
A:Title: Polymorphism of the 3' open reading frame of the virus associated with the acqu
A:Reference number: S03244; MUID:86067228
A:Accession: S03244
A:Molecule type: DNA
A:Residues: 1-206 <RAT>
A:Cross-references: EMBL:X03187
A:Note: The in-frame stop codon at residue 124, shown as 'X', may be suppressed
C:Genetics:
A:Gene: nef; 3'-orf; orf-F
C:Superfamily: AIDS nef protein

Query Match 66.1%; Score 1115; DB 2; Length 206;
Best Local Similarity 98.5%; Pred. No. 3.2e-85;
Matches 203; Conservative 2; Mismatches 1; Indels 0; Gaps 0;

QY 1 MGGKSSVVGWPTVRRMRRAEPAADGGAASRDLEKHGAISSNTAATNAACAWLEA 60
|||||
DB 1 MGGKSSVVGWPTVRRMRRAEPAADGGAASRDLEKHGAISSNTAATNAACAWLEA 60
|||||
QY 61 QEEVGFPVTPQVPLRPMTYKAADVLSHFLKEKGLGLIHSQRQDILDLYHTQGY 120
|||||
DB 61 QEEVGFPVTPQVPLRPMTYKAADVLSHFLKEKGLGLIHSQRQDILDLYHTQGY 120
|||||
QY 121 FPDQWNTYTPGQVRYPLTFGWCYKLVPEPDKVEEANKGENTSLHPVSLHGMDDPEREV 180
|||||
```

```
DB 121 FPDQWNTYTPGQVRYPLTFGWCYKLVPEPDKVEEANKGENTSLHPVSLHGMDDPEREV 180
QY 181 LEWRFDLSRLAFHHVARELHPEYFKNC 206
|||||
DB 181 LEWRFDLSRLAFHHVARELHPEYFKNC 206
|||||

RESULT 4
ASLJVL
nef protein - human immunodeficiency virus type 1 (isolate LV)
N:Alternate names: 3'-orf protein; orf-F protein
C:Species: human immunodeficiency virus type 1, HIV-1
C:Date: 17-May-1985 #sequence_revision 17-May-1985 #text_change 16-Jul-1999
C:Accession: A04007
R:Muesing, M.A.; Smith, D.H.; Cabradilla, C.D.; Benton, C.V.; Lasky, L.A.; Capon, D.J
Nature 313, 450-458, 1985
A:Title: Nucleic acid structure and expression of the human AIDS/lymphadenopathy retr
A:Reference number: A93355; MUID:85111157
A:Accession: A04007
A:Molecule type: DNA
A:Residues: 1-206 <MUE>
A:Cross-references: GB:K02083; NID:G555008; PIDN:AAB59874.1; PID:G328560
C:Genetics:
A:Gene: nef; 3'-orf; orf-F
C:Superfamily: AIDS nef protein
C:Keywords: AIDS; immunodeficiency

Query Match 66.0%; Score 1114; DB 1; Length 206;
Best Local Similarity 97.6%; Pred. No. 3.8e-85;
Matches 201; Conservative 3; Mismatches 2; Indels 0; Gaps 0;

QY 1 MGGKSSVVGWPTVRRMRRAEPAADGGAASRDLEKHGAISSNTAATNAACAWLEA 60
|||||
DB 1 MGGKSSVVGWPAVRMRRAEPAADGGAASRDLEKHGAISSNTAATNAACAWLEA 60
|||||
QY 61 QEEVGFPVTPQVPLRPMTYKAADVLSHFLKEKGLGLIHSQRQDILDLYHTQGY 120
|||||
DB 61 QEEVGFPVTPQVPLRPMTYKAADVLSHFLKEKGLGLIHSQRQDILDLYHTQGY 120
|||||
QY 121 FPDQWNTYTPGQVRYPLTFGWCYKLVPEPDKVEEANKGENTSLHPVSLHGMDDPEREV 180
|||||
DB 121 FPDQWNTYTPGQVRYPLTFGWCYKLVPEPDKVEEANKGENTSLHPVSLHGMDDPEREV 180
|||||
QY 181 LEWRFDLSRLAFHHVARELHPEYFKNC 206
|||||
DB 181 LEWRFDLSRLAFHHVARELHPEYFKNC 206
|||||

RESULT 5
ASLJ12
nef protein - human immunodeficiency virus type 1 (isolate HTLV-III, 12)
N:Alternate names: 3'-orf protein; orf-F protein
C:Species: human immunodeficiency virus type 1, HIV-1
A:Note: host Homo sapiens (man)
C:Date: 04-Dec-1986 #sequence_revision 31-Dec-1991 #text_change 16-Jul-1999
C:Accession: A04006
R:Arya, S.K.; Gallo, R.C.
Proc. Natl. Acad. Sci. U.S.A. 83, 2209-2213, 1986
A:Title: Three novel genes of human T-lymphotropic virus type III: immune reactivity
A:Reference number: A94093; MUID:86177573
A:Accession: A04006
A:Molecule type: DNA
A:Residues: 1-206 <ARY>
A:Cross-references: EMBL:M11840; NID:G328453; PIDN:AAA5001.1; PID:G328458
C:Genetics:
A:Gene: nef; 3'-orf; orf-F
C:Superfamily: AIDS nef protein
C:Keywords: AIDS; immunodeficiency

Query Match 65.6%; Score 1108; DB 1; Length 206;
Best Local Similarity 97.1%; Pred. No. 1.2e-84;
```

GenCore version 4.5  
Copyright (c) 1993 - 2000 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: August 26, 2002, 08:08:31 ; Search time 50.45 Seconds  
(Without alignments)  
575.203 Million cell updates/sec

Title: US-09-509-239-13  
Perfect score: 1688  
Sequence: 1 MGGKWSKVVGWPTVRM.....QSRCDPTGPKTSGHHHHH 302

Scoring table: BLOSUM62  
Gapop 10.0 , Gapext 0.5

Searched: 283138 seqs, 96089334 residues

Total number of hits satisfying chosen parameters: 283138

Minimum DB seq length: 0  
Maximum DB seq length: 2000000000  
Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

Database : PIR\_71:\*  
1: pir1:\*  
2: pir2:\*  
3: pir3:\*  
4: pir4:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

## SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	1130	66.9	206	1 ASLJFV	nef protein - huma
2	1115	66.1	205	2 S43467	nef protein - huma
3	1115	66.1	206	2 S03244	nef protein (clone
4	1114	66.0	206	1 ASLJVL	nef protein - huma
5	1108	65.6	206	1 ASLJL2	nef protein - huma
6	1097	65.0	206	2 JC5400	nef protein - huma
7	1096	64.9	206	2 S03245	nef protein (clone
8	1096	64.9	206	2 S33986	nef protein - huma
9	1081.5	64.1	205	1 ASLJH3	nef protein - huma
10	1019	60.4	204	2 S24985	nef protein - huma
11	1019	60.4	218	1 ASLJBR	nef protein - huma
12	1007	59.7	206	2 S25937	nef protein - huma
13	997	59.1	210	1 ASLJ02	nef protein - huma
14	991.5	58.7	214	1 I44001	nef protein - huma
15	987	58.5	182	2 S03247	nef protein (clone
16	987	58.5	204	2 S03246	nef protein - huma
17	911	54.0	212	1 QQLJ2R	nef protein (clone
18	903.5	53.5	207	1 QQLJND	nef protein - huma
19	837	49.6	205	1 B44963	nef protein - huma
20	834.5	49.4	209	2 T01673	nef protein - huma
21	832	49.3	205	1 ASLJIK	nef protein - huma
22	486	28.8	86	2 S33982	nef protein - simi
23	478	28.3	95	1 TNLJ12	trans-activating t
24	470	27.8	86	2 A25700	trans-activating t
25	443	26.2	226	2 S46353	trans-activating t
26	439.5	26.0	97	2 S54385	nef protein - simi
27	437	25.9	101	2 T09446	nef protein - huma
28	434	25.7	101	1 E44001	tat protein - huma
29	429.5	25.4	309	2 S07993	trans-activating t

## ALIGNMENTS

## RESULT 1

ASLJFV

nef protein - human immunodeficiency virus type 1 (isolate LAV-la)

N:Alternate names: 3'-orf protein; orf-F protein

C:Species: human immunodeficiency virus type 1, HIV-1

A:Note: host Homo sapiens (man)

C:Date: 17-May-1985 #sequence\_revision 17-May-1985 #text\_change 16-Jul-1999

C:Accession: A04008; S14609

R:Wain-Hobson, S.; Sonigo, P.; Danos, O.; Cole, S.; Alizon, M.

Cell 40, 9-17, 1985

A:Title: Nucleotide sequence of the AIDS virus, LAV.

A:Reference number: A30866; MUID:85099333

A:Accession: A04008

A:Molecule type: DNA

A:Residues: 1-206 &lt;WAT&gt;

A:Cross-references: GB:K02013; NID:g326417; PIDN:AAB59752.1; PID:g326425

A:Experimental source: isolate LAV-la

R:Ciccarelli, R.B.

Submitted to the EMBL Data Library, March 1991

A:Reference number: S14607

A:Accession: S14609

A&gt;Status: preliminary

A:Molecule type: DNA

A:Residues: 1-206 &lt;CIC&gt;

A:Cross-references: EMBL:X58780; NID:g60113; PIDN:CAA41585.1; PID:g60114

C:Genetics:

A:Gene: nef; 3'-orf; orf-F

C:Superfamily: AIDS nef protein

C:Keywords: AIDS; immunodeficiency; phosphoprotein

## Query Match

66.9% Score 1130; DB 1; Length 206;

Best Local Similarity 100.0%; Pred. No. 1.8e-86;

Matches 206; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 MGGKWSKVVGWPTVRMRRAEPAADGVGAASRDLEKHGAITSNTAATNAACAWLEA 60

Db 1 MGGKWSKVVGWPTVRMRRAEPAADGVGAASRDLEKHGAITSNTAATNAACAWLEA 60

Qy 61 QEEEVGFPVTPQVPLRPMYKAAVDLSHFLKEKGGLEGLIHSQRQDILDWVHTQGY 120

Db 61 QEEEVGFPVTPQVPLRPMYKAAVDLSHFLKEKGGLEGLIHSQRQDILDWVHTQGY 120

Qy 121 FPDQNYTPGPGVRYPLTFGCYKLVPPDPKVEANKGENTSLHPVSLHGMDPPEV 180

Db 121 FPDQNYTPGPGVRYPLTFGCYKLVPPDPKVEANKGENTSLHPVSLHGMDPPEV 180

Qy 181 LEWRFDLSLAFHHVARELHPEYFKNC 206

Db 181 LEWRFDLSLAFHHVARELHPEYFKNC 206



Matches 200; Conservative 4; Mismatches 2; Indels 0; Gaps 0;

QY 1 MGGKWSKSSVVGWPTVRRMRRAEPAADGVGAASRDLEKKGAISSNTAATNAACAWLEA 60  
 DB 1 MGGKWSKSSVVGWPAVRMRRAEPAADGVGAASRDLEKKGAISSNTAATNAACAWLEA 60

QY 61 QEEVEVGFPTVQVPLRPMTYKAAVDLSHFLKEKGGLEGLIHSQRRODILDLYIHTQGY 120  
 DB 61 QEEKVGFPVTPQVPLRPMTYKAAVDLSHFLKEKGGLEGLIHSQRRODILDLYIHTQGY 120

QY 121 FPDQNYTPGCVRYPLTFGWCYKLVPEPKVEANKGENTSLHHPVSLHGMDPPER 180  
 DB 121 FPDQNYTPGCVRYPLTFGWCYKLVPEPKVEANKGENTSLHHPVSLHGMDPPER 180

QY 181 LEWRDLSRLAFHHVARELHPEYFKNK 206  
 DB 181 LEWRDLSRLAFHHVARELHPEYFKNK 206

RESULT 6  
 JC5400  
 nef protein - Human immunodeficiency virus type 1, HIV-1  
 C:Species: Human immunodeficiency virus type 1, HIV-1  
 C:Date: 07-Jul-1997 #sequence\_revision 18-Jul-1997 #text\_change 31-Oct-1997  
 C:Accession: JC5400  
 R:Macreadie, I.G.; Lowe, M.G.; Curtain, C.C.; Hewish, D.; Azad, A.A.  
 Biochem. Biophys. Res. Commun. 232, 707-711, 1997  
 A:Title: Cytotoxicity resulting from addition of HIV-1 Nef N-terminal peptides to yeast  
 A:Reference number: JC5400; MUID:97271389  
 A:Accession: JC5400  
 A:Molecule type: protein  
 A:Residues: 1-206 <MAC>  
 C:Comment: The amino-terminal part possesses membrane-perturbing and fusogenic activities.  
 C:Superfamily: AIDS nef protein  
 F:2-206/Product: nef protein #status predicted <MAT>

Query Match 65.0%; Score 1097; DB 2; Length 206;  
 Best Local Similarity 97.1%; Pred. No. 9.9e-84;  
 Matches 200; Conservative 1; Mismatches 5; Indels 0; Gaps 0;

QY 1 MGGKWSKSSVVGWPTVRRMRRAEPAADGVGAASRDLEKKGAISSNTAATNAACAWLEA 60  
 DB 1 MGGKWSKSSVVGWPAVRMRRAEPAADGVGAASRDLEKKGAISSNTAATNAACAWLEA 60

QY 61 QEEVEVGFPTVQVPLRPMTYKAAVDLSHFLKEKGGLEGLIHSQRRODILDLYIHTQGY 120  
 DB 61 QEEVEVGFPTVQVPLRPMTYKAAVDLSHFLKEKGGLEGLIHSQRRODILDLYIHTQGY 120

QY 121 FPDQNYTPGCVRYPLTFGWCYKLVPEPKVEANKGENTSLHHPVSLHGMDPPER 180  
 DB 121 FPDQNYTPGCVRYPLTFGWCYKLVPEPKVEANKGENTSLHHPVSLHGMDPPER 180

QY 181 LEWRDLSRLAFHHVARELHPEYFKNK 206  
 DB 181 LEWRDLSRLAFHHVARELHPEYFKNK 206

RESULT 7  
 S03245  
 nef protein (clone HXB3) - human immunodeficiency virus type 1  
 N:Alternate names: 3'-orf protein  
 C:Species: human immunodeficiency virus type 1, HIV-1  
 C:Date: 28-Feb-1990 #sequence\_revision 28-Feb-1990 #text\_change 20-Sep-1999  
 C:Accession: S03245  
 R:Ratner, L.; Starcich, B.; Josephs, S.F.; Hahn, B.H.; Reddy, K.J.; Petteway  
 Nucleic Acids Res. 13, 8219-8229, 1985  
 A:Title: Polymorphism of the 3' open reading frame of the virus associated with the acq  
 A:Reference number: S03244; MUID:86067228  
 A:Accession: S03245  
 A:Molecule type: DNA  
 A:Residues: 1-206 <RAT>

A:Cross-references: EMBL:X03188; NID:g61556; PIDN:CAA26947.1  
 C:Genetics:  
 A:Gene: nef; 3'-orf; Orf-F  
 C:Superfamily: AIDS nef protein

Query Match 64.9%; Score 1096; DB 2; Length 206;  
 Best Local Similarity 96.6%; Pred. No. 1.2e-83;  
 Matches 199; Conservative 4; Mismatches 3; Indels 0; Gaps 0;

QY 1 MGGKWSKSSVVGWPTVRRMRRAEPAADGVGAASRDLEKKGAISSNTAATNAACAWLEA 60  
 DB 1 MGGKWSKSSVVGWPAVRMRRAEPAADGVGAASRDLEKKGAISSNTAATNAACAWLEA 60

QY 61 QEEVEVGFPTVQVPLRPMTYKAAVDLSHFLKEKGGLEGLIHSQRRODILDLYIHTQGY 120  
 DB 61 QEEKVGFPVTPQVPLRPMTYKAAVDLSHFLKEKGGLEGLIHSQRRODILDLYIHTQGY 120

QY 121 FPDQNYTPGCVRYPLTFGWCYKLVPEPKVEANKGENTSLHHPVSLHGMDPPER 180  
 DB 121 FPDQNYTPGCVRYPLTFGWCYKLVPEPKVEANKGENTSLHHPVSLHGMDPPER 180

QY 181 LEWRDLSRLAFHHVARELHPEYFKNK 206  
 DB 181 LEWRDLSRLAFHHVARELHPEYFKNK 206

RESULT 8  
 S33986  
 nef protein - human immunodeficiency virus type 1  
 C:Species: human immunodeficiency virus type 1, HIV-1  
 C:Date: 19-Mar-1997 #sequence\_revision 19-Mar-1997 #text\_change 20-Sep-1999  
 C:Accession: S33986  
 R:Carlini, F.  
 submitted to the EMBL Data Library, November 1991  
 A:Reference number: S33979  
 A:Accession: S33986  
 A:Status: preliminary  
 A:Molecule type: mRNA  
 A:Residues: 1-206 <CAR>  
 A:Cross-references: EMBL:Z11530; NID:g60192; PIDN:CAA77629.1; PID:g60200  
 C:Superfamily: AIDS nef protein

Query Match 64.9%; Score 1096; DB 2; Length 206;  
 Best Local Similarity 96.6%; Pred. No. 1.2e-83;  
 Matches 199; Conservative 3; Mismatches 4; Indels 0; Gaps 0;

QY 1 MGGKWSKSSVVGWPTVRRMRRAEPAADGVGAASRDLEKKGAISSNTAATNAACAWLEA 60  
 DB 1 MGGKWSKSSVVGWPAVRMRRAEPAADGVGAASRDLEKKGAISSNTAATNAACAWLEA 60

QY 61 QEEVEVGFPTVQVPLRPMTYKAAVDLSHFLKEKGGLEGLIHSQRRODILDLYIHTQGY 120  
 DB 61 QEEVEVGFPTVQVPLRPMTYKAAVDLSHFLKEKGGLEGLIHSQRRODILDLYIHTQGY 120

QY 121 FPDQNYTPGCVRYPLTFGWCYKLVPEPKVEANKGENTSLHHPVSLHGMDPPER 180  
 DB 121 FPDQNYTPGCVRYPLTFGWCYKLVPEPKVEANKGENTSLHHPVSLHGMDPPER 180

QY 181 LEWRDLSRLAFHHVARELHPEYFKNK 206  
 DB 181 LEWRDLSRLAFHHVARELHPEYFKNK 206

RESULT 9  
 ASLJH3  
 nef protein - human immunodeficiency virus type 1 (isolate HTLV-III, BH10)  
 N:Alternate names: 3'-orf protein; Orf-F protein  
 C:Species: human immunodeficiency virus type 1, HIV-1  
 A:Note: host Homo sapiens (man)  
 C:Date: 17-May-1985 #sequence\_revision 17-May-1985 #text\_change 31-Jan-1997  
 C:Accession: A04005

Streicher,

R:Ratner, L.; Haseltine, W.; Patarca, R.; Livak, K.J.; Starcich, B.; Josephs, S.F.; Dora  
nberger, J.A.; Papas, T.S.; Chirayeb, J.; Chang, N.T.; Gallo, R.C.; Wong-Staal, F.  
Nature 313, 277-284, 1985

A:Title: Complete nucleotide sequence of the AIDS virus, HTLV-III.

A:Reference number: A93353; MUID:85111123

A:Accession: A04005

A:Molecule type: DNA

A:Residues: 1-205 <RAT>

C:Genetics:

A:Gene: nef; 3'-orf; Orf-F

C:Superfamily: AIDS nef protein

C:Keywords: AIDS; immunodeficiency

```
Query Match          64.1%; Score 1081.5; DB 1; Length 205;
Best Local Similarity 96.1%; Pred. No. 1.9e-82;
Matches 198; Conservative 4; Mismatches 3; Indels 1; Gaps 1;

QY 1 MGGKWSKSSVVGWPTVRMRRAEPAADGVGAASRDLEKHGAITSSNTAATNAACAWLEA 60
DB 1 MGGKWSKSSVVGWPAVRMRRAEPAADGVGAASRDLEKHGAITSSNTAANADCAWLEA 60

QY 61 QEEVEGFPVTPQVPLRPMTYKAAVDLSHFLKEKGLGLIHSQRQDILDWYHTQGY 120
DB 61 QEEVEGFPVTPQVPLRPMTYKAAVDLSHFLKEKGLGLIHSQRQDILDWYHTQGY 120

QY 121 FPDQNYTPGCGVRYPLTFGCWCKLVPEPKVEANGKENTSLHPVSLHGMDPPERV 180
DB 121 FPDQNYTPGCGVRYPLTFGCWCKLVPEPKVEANGKENTSLHPVSLHGMDPPERV 180

QY 181 LEWRDSDSLAFHHMARELHPEYFKNC 206
DB 181 LEWRDSDSLAFHHMARELHPEYFKNC 205
```

```
RESULT 10
S24985
nef protein - human immunodeficiency virus type 1
C:Species: human immunodeficiency virus type 1, HIV-1
C:Date: 20-Feb-1995 #sequence_revision 20-Feb-1995 #text_change 20-Sep-1999
C:Accession: S24985
R:Harris, M.; Hislop, S.; Patsilinos, P.; Neil, J.C.
A:Description: In vivo derived HIV-1 NEF genes are heterogeneous and lack detectable nuc
A:Reference number: S24985
A:Accession: S24985
A>Status: preliminary
A:Molecule type: DNA
A:Residues: 1-204 <HAR>
A:Cross-references: EMBL:X63044; NID:g60123; PIDN:CAA44770.1; PID:g60124
C:Superfamily: AIDS nef protein
```

```
Query Match          60.4%; Score 1019; DB 2; Length 204;
Best Local Similarity 89.8%; Pred. No. 2.9e-77;
Matches 185; Conservative 10; Mismatches 9; Indels 2; Gaps 1;

QY 1 MGGKWSKSSVVGWPTVRMRRAEPAADGVGAASRDLEKHGAITSSNTAATNAACAWLEA 60
DB 1 MGGKWSKSSVVGWPTVRMRRAEPAADGVGAASRDLEKHGAITSSNTAATNAACAWLEA 58

QY 61 QEEVEGFPVTPQVPLRPMTYKAAVDLSHFLKEKGLGLIHSQRQDILDWYHTQGY 120
DB 59 QEEVEGFPVTPQVPLRPMTYKAAVDLSHFLKEKGLGLIHSQRQDILDWYHTQGY 118

QY 121 FPDQNYTPGCGVRYPLTFGCWCKLVPEPKVEANGKENTSLHPVSLHGMDPPERV 180
DB 119 FPDQNYTPGCGVRYPLTFGCWCKLVPEPKVEANGKENTSLHPVSLHGMDPPERV 178

QY 181 LEWRDSDSLAFHHMARELHPEYFKNC 206
DB 179 LEWRDSDSLAFHHMARELHPEYFKNC 204
```

## RESULT 11

ASLJBR

nef protein - human immunodeficiency virus type 1 (isolate BR)

N:Alternate names: 3'-orf protein; Orf-F protein

C:Species: human immunodeficiency virus type 1, HIV-1

A:Note: host Homo sapiens (man)

C:Date: 31-Dec-1989 #sequence\_revision 31-Dec-1989 #text\_change 16-Jul-1999

C:Accession: D31667; S21993; S21995; S21997; S21999; S21991

R:Anand, R.; Thayer, R.; Srinivasan, A.; Nayyar, S.; Gardner, M.; Luciw, P.; Dandeka

Virolgy 168, 79-89, 1989

A:Title: Biological and molecular characterization of human immunodeficiency virus (

A:Reference number: A94389; MUID:89085613

A:Accession: D31667

A:Molecule type: DNA

A:Residues: 1-218 <ANA>

A:Cross-references: GB:M21098; NID:g326426; PIDN:AAA44222.1; PID:g326431

R:Steuler, H.; Storch-Hagenlocher, B.; Wildemann, B.; Hacke, W.

submitted to the EMBL Data Library, July 1991

A:Description: Distinct populations of HIV-1 in blood and cerebrospinal fluid as det

A:Reference number: S21990

A:Accession: S21993

A:Molecule type: DNA

A:Residues: 1-7 <STE>

A:Cross-references: EMBL:X61358; NID:g60177; PIDN:CAA43629.1; PID:g584027; EMBL:X613

A:Accession: S21995

A:Molecule type: DNA

A:Residues: 1-7 <ST2>

A:Cross-references: EMBL:X61355; NID:g60179; PIDN:CAA43623.1; PID:g584028; EMBL:X613

A:Accession: S21997

A:Molecule type: DNA

A:Residues: 1-7 <ST3>

A:Cross-references: EMBL:X61356; NID:g60181; PIDN:CAA43625.1; PID:g1129140; EMBL:X61

A:Accession: S21999

A:Molecule type: DNA

A:Residues: 1-7 <ST4>

A:Cross-references: EMBL:X61359; NID:g60182; PIDN:CAA43631.1; PID:g584030

C:Genetics:

A:Gene: nef; 3'-orf; Orf-F

C:Superfamily: AIDS nef protein

C:Keywords: AIDS; immunodeficiency

## Query Match

Best Local Similarity 84.1%; Score 1019; DB 1; Length 218;

Matches 185; Conservative 12; Mismatches 7; Indels 16; Gaps 2;

QY 1 MGGKWSKSSVVGWPTVRMR-----RAEPAADGVGAASRDLEKHGAITSS 46

DB 1 MGGKWSK--MAGWSTVRMRRAEPAEPAADGVGAASRDLEKHGAITSS 58

QY 47 NTAATNAACAWLEAQEEEEVGFPVTPQVPLRPMTYKAAVDLSHFLKEKGLGLIHSQR 106

DB 59 NTAATNAACAWLEAQEEDEVEGFPVTPQVPLRPMTYKAAVDLSHFLKEKGLGLIHSQR 118

QY 107 QDILDWYHTQGYFPDQNYTPGCGVRYPLTFGCWCKLVPEPKVEANGKENTSLH 166

DB 119 QDILDWYHTQGYFPDQNYTPGCGVRYPLTFGCWCKLVPEPKVEANGKENTSLH 178

QY 167 PVSLHGMDPPERVLEWFRDLSRLAFHHMARELHPEYFKNC 206

DB 179 PMSQHGMDPPERVLEWFRDLSRLAFHHMARELHPEYFKNC 218

## RESULT 12

S25937

nef protein - human immunodeficiency virus type 1

C:Species: human immunodeficiency virus type 1, HIV-1

C:Date: 25-Feb-1994 #sequence\_revision 01-Dec-1995 #text\_change 20-Sep-1999

C:Accession: S25937

R:Guo, H.G.; Chernann, J.C.; Waters, D.; Hall, L.; Louie, A.; Gallo, R.C.;

Nature 349, 745-746, 1991

A:Title: Sequence analysis of original HIV-1.

A:Reference number: S25937; MUID:91156044  
 A:Accession: S25937  
 A:Status: preliminary; nucleic acid sequence not shown; translation not shown  
 A:Molecule type: nucleic acid  
 A:Residues: 1-206 <GUC>  
 A:Cross-references: EMBL:X57465; NID:g60217; PIDN:CAA40702.1; PID:g60218  
 A:Note: the nucleotide sequence was submitted to the EMBL Data Library, March 1991  
 C:Superfamily: AIDS nef protein

Query Match 59.7%; Score 1007; DB 2; Length 206;  
 Best Local Similarity 85.9%; Pred. No. 2.9e-76;  
 Matches 177; Conservative 18; Mismatches 11; Indels 0; Gaps 0;

QY 1 MGCKWKSXSVVGVPTVRMRRAEPAADGVGAASRDLEKKGAISSNTAATNAACAWLEA 60  
 DB 1 MGCKWKSXSVVGVPTVRMRRAEPAADGVGAASRDLEKKGAISSNTAATNAACAWLEA 60  
 QY 61 QEEVEYFPPVTPQVPLRPMTYKAADVLSHFLKKGGLGGLIHSQRQDILDLWIYHTQGY 120  
 DB 61 QEEVEYFPPVTPQVPLRPMTYKAADVLSHFLKKGGLGGLIHSQRQDILDLWIYHTQGY 120  
 QY 121 FPDWQNTGPGVRYPLTFCWCYKLVPEPDKVEEANKGENTSLHPVSLHGMDPEREV 180  
 DB 121 FPDWQNTGPGVRYPLTFCWCYKLVPEPDKVEEANKGENTSLHPVSLHGMDPEREV 180  
 QY 181 LEWRFSRLAFHVAHRELHPEYKNC 206  
 DB 181 LEWRFSRLAFHVAHRELHPEYKNC 206

RESULT 13  
 ASLJO2  
 nef protein - human immunodeficiency virus type 1 (isolate ARV-2)  
 N:Alternate names: 3'-orf protein; orf-F protein  
 C:Species: human immunodeficiency virus type 1, HIV-1  
 A:Note: host Homo sapiens (man)  
 C:Date: 17-May-1985 #sequence\_revision 17-May-1985 #text\_change 16-Jul-1999  
 C:Accession: A04009  
 R:Sanchez-Pescador, R.; Power, M.D.; Barr, P.J.; Steimer, K.S.; Stempien, M.M.; Brown SH  
 Science 227, 484-492, 1985  
 A:Title: Nucleotide sequence and expression of an AIDS-associated retrovirus (ARV-2).  
 A:Reference number: A04003; MUID:85090453  
 A:Accession: A04009  
 A:Molecule type: DNA  
 A:Residues: 1-210 <SAN>  
 A:Cross-references: GB:K02007; NID:g328658; PIDN:AAB59883.1; PID:g328667  
 C:Genetics:  
 A:Gene: nef; 3'-orf; orf-F  
 C:Superfamily: AIDS nef protein  
 C:Keywords: AIDS; Immunodeficiency

Query Match 59.1%; Score 997; DB 1; Length 210;  
 Best Local Similarity 85.2%; Pred. No. 2e-75;  
 Matches 179; Conservative 16; Mismatches 11; Indels 4; Gaps 1;

QY 1 MGCKWKSXSVVGVPTVRMRRAEPAADGVGAASRDLEKKGAISSNTAATNAACAA 56  
 DB 1 MGCKWKSXSVVGVPTVRMRRAEPAADGVGAASRDLEKKGAISSNTAATNAACAA 56  
 QY 57 WLEAQEEVEYFPPVTPQVPLRPMTYKAADVLSHFLKKGGLGGLIHSQRQDILDLWIYH 116  
 DB 57 WLEAQEEVEYFPPVTPQVPLRPMTYKAADVLSHFLKKGGLGGLIHSQRQDILDLWIYH 116  
 QY 117 TQGYFPDQWNTGPGVRYPLTFCWCYKLVPEPDKVEEANKGENTSLHPVSLHGMDP 176  
 DB 117 TQGYFPDQWNTGPGVRYPLTFCWCYKLVPEPDKVEEANKGENTSLHPVSLHGMDP 176  
 QY 177 EREVLEWRFSRLAFHVAHRELHPEYKNC 206  
 DB 181 EREVLEWRFSRLAFHVAHRELHPEYKNC 210

RESULT 14  
 I44001  
 nef protein - human immunodeficiency virus type 1 (strain YU-2)  
 N:Alternate names: 3'-orf protein; orf-F protein  
 C:Species: human immunodeficiency virus type 1, HIV-1  
 A:Note: host Homo sapiens (man)  
 C:Date: 31-Dec-1993 #sequence\_revision 31-Dec-1993 #text\_change 23-Feb-1997  
 C:Accession: I44001  
 R:Li, Y.; Hui, H.; Burgess, C.J.; Price, R.W.; Sharp, P.M.; Hahn, B.H.; Shaw, G.M.  
 J. Virol. 66, 6587-6600, 1992  
 A:Title: Complete nucleotide sequence, genome organization, and biological properties  
 A:Reference number: A44001; MUID:93021387  
 A:Accession: I44001  
 A:Molecule type: DNA  
 A:Residues: 1-214 <LIY>  
 A:Cross-references: GB:M93258  
 C:Genetics:  
 A:Gene: nef; 3'-orf; orf-F  
 C:Superfamily: AIDS nef protein  
 C:Keywords: AIDS; Immunodeficiency

Query Match 58.7%; Score 991.5; DB 1; Length 214;  
 Best Local Similarity 84.2%; Pred. No. 5.9e-75;  
 Matches 181; Conservative 12; Mismatches 11; Indels 11; Gaps 2;  
 QY 1 MGCKWKSXSVVGVPTVRRAEPAADGVGAASRDLEKKGAISSNTAA 50  
 DB 1 MGCKWKSXSVVGVPTVRRAEPAADGVGAASRDLEKKGAISSNTAA 60  
 QY 51 TNAACAWLEAQEEVEYFPPVTPQVPLRPMTYKAADVLSHFLKKGGLGGLIHSQRQDIL 110  
 DB 61 TNADCAWLEAQEEVEYFPPVTPQVPLRPMTYKAADVLSHFLKKGGLGGLIHSQRQDIL 120  
 QY 111 DLWIYHTQGYFPDQWNTGPGVRYPLTFCWCYKLVPEPDKVEEANKGENTSLHPVSL 170  
 DB 121 DLWIYHTQGYFPDQWNTGPGVRYPLTFCWCYKLVPEPDKVEEANKGENTSLHPVSL 179  
 QY 171 HGMDPEREVLEWRFSRLAFHVAHRELHPEYKNC 205  
 DB 180 HGMDPEREVLEWRFSRLAFHVAHRELHPEYKNC 214

RESULT 15  
 S03247  
 nef protein (clone C15) - human immunodeficiency virus type 1 (fragment)  
 N:Alternate names: 3'-orf protein  
 C:Species: human immunodeficiency virus type 1, HIV-1  
 C:Date: 28-Feb-1990 #sequence\_revision 28-Feb-1990 #text\_change 20-Jun-2000  
 C:Accession: S03247  
 R:Ratner, L.; Starcich, B.; Josephs, S.F.; Hahn, B.H.; Reddy, E.P.; Livak, K.J.; Pett  
 Nucleic Acids Res. 13, 8219-8229, 1985  
 A:Title: Polymorphism of the 3' open reading frame of the virus associated with the a  
 A:Reference number: S03244; MUID:86067228  
 A:Accession: S03247  
 A:Molecule type: mRNA  
 A:Residues: 1-192 <RAT>  
 A:Cross-references: EMBL:X03189; NID:g61552; PIDN:CAA26948.1; PID:g1335562  
 C:Genetics:  
 A:Gene: nef; 3'-orf; orf-F  
 C:Superfamily: AIDS nef protein

Query Match 58.5%; Score 987; DB 2; Length 182;  
 Best Local Similarity 98.4%; Pred. No. 1.2e-74;  
 Matches 179; Conservative 1; Mismatches 2; Indels 0; Gaps 0;

QY 25 PAADGVGAASRDLEKKGAISSNTAATNAACAWLEAQEEVEYFPPVTPQVPLRPMTYKAA 84  
 DB 1 PAADGVGAASRDLEKKGAISSNTAATNAACAWLEAQEEVEYFPPVTPQVPLRPMTYKAA 60  
 QY 85 VDLSHFLKKGGLGGLIHSQRQDILDLWIYHTQGYFPDQWNTGPGVRYPLTFCWCYK 144

Db 61 VDLSHFLKEKGGLEGLIHSQRQDILDWYHTQGYFPDQWNYTPGCVRYPLTFGWCYK 120  
QY 145 LVPVDPKVEEANKGENTSLHVPVSLHGMDPEREVLEWRDLSRLAFHHVARELHPEYFK 204  
Db 121 LVPVDPKVEEANKGENTSLHVPVSLHGMDPEREVLEWRDLSRLAFHHVARELHPEYFK 180  
QY 205 NC 206  
Db 181 NC 182

RESULT 16  
S03246  
nef protein (clone HAT3) - human immunodeficiency virus type 1  
N:Alternate names: 3'-orf protein  
C:Species: human immunodeficiency virus type 1, HIV-1  
C:Date: 28-Feb-1990 #sequence\_revision 28-Feb-1990 #text\_change 20-Sep-1999  
C:Accession: S03246  
R:Ratner, L.; Starcich, B.; Josephs, S.F.; Hahn, B.H.; Reddy, E.P.; Livak, K.J.; Pettawa  
Nucleic Acids Res. 13, 8219-8229, 1985  
A:Title: Polymorphism of the 3' open reading frame of the virus associated with the acqu  
A:Reference number: S03244; MUID:86067228  
A:Accession: S03246  
A:Molecule type: DNA  
A:Residues: 1-204 <RAT>  
A:Cross-references: EMBL:X03190; NID:g61550; PIDN:CAA26949.1; PID:g61551  
A:Note: The authors translated the codon AGT for residue 11 as Gly  
C:Genetics:  
A:Gene: nef; 3'-orf; orf-F  
C:Superfamily: AIDS nef protein

Query Match 58.5%; Score 987; DB 2; Length 204;  
Best Local Similarity 84.5%; Pred. No. 1.3e-74; Mismatches 14; Indels 2; Gaps 1;  
Matches 174; Conservative 16;  
QY 1 MGGKWSKSSVVGWPTVRRMRRAEPAADGVGAASRDLEKHGAISSNTAATNAACAWLEA 60  
Db 1 MGGKWSKSKMGPVAVRERMQRAEPAADGVGAASRDLEKHGAISSNT--NNAACTWLEA 58  
QY 61 QEEVEGFPVTPQVPLRPMTYKAADVLSHFLKEKGGLEGLIHSQRQDILDWYHTQY 120  
Db 59 QEEVEGFPVTPQVPLRPMTYKAADVLSHFLKEKGGLEGLIHSQRQDILDWYHTQY 118  
QY 121 FPDQWNYTPGCVRYPLTFGWCYKLVPEPDKVEEANKGENTSLHVPVSLHGMDPEREV 180  
Db 119 FPDQWNYTPGCVRYPLTFGWCYKLVPEPDKVEEANKGENTSLHVPVSLHGMDPEREV 178  
QY 181 LEWRDLSRLAFHHVARELHPEYFKNC 206  
Db 179 LVNKFDSRLAFHHVARELHPEYFKNC 204

RESULT 17  
QOLJZR  
nef protein - human immunodeficiency virus Zr-6  
N:Alternate names: 3'-orf protein; orf-F protein  
C:Species: human immunodeficiency virus Zr-6  
C:Date: 30-Sep-1987 #sequence\_revision 30-Sep-1987 #text\_change 16-Jul-1999  
C:Accession: F26192  
R:Srinivasan, A.; Anand, R.; York, D.; Ranganathan, P.; Feorino, P.; Schochetman, G.; Cu  
Gene 52, 71-82, 1987  
A:Title: Molecular characterization of human immunodeficiency virus from Zaire: nucleoti  
A:Reference number: A26192; MUID:87248097  
A:Accession: F26192  
A:Molecule type: DNA  
A:Residues: 1-212 <SRI>  
A:Cross-references: GB:K03458; GB:M16322; NID:g329398; PIDN:AAA45381.1; PID:g329404  
C:Genetics:  
A:Gene: nef; 3'-orf; orf-F  
C:Superfamily: AIDS nef protein  
C:Keywords: AIDS; immunodeficiency

Query Match 54.0%; Score 911; DB 1; Length 212;  
Best Local Similarity 77.4%; Pred. No. 2.8e-68;  
Matches 164; Conservative 23; Mismatches 19; Indels 6; Gaps 2;

QY 1 MGGKWSKSSVVGWPTVRRMRRAEPAADGVGAASRDLEKHGAISSNTAATNAAC 55  
Db 1 MGGKWSKSSVVGWPTVRRMRRAEPAADGVGAASRDLEKHGAISSNTAATNAAC 60  
QY 56 AWLEAQEE-BEVEGFPVTPQVPLRPMTYKAADVLSHFLKEKGGLEGLIHSQRQDILDWY 114  
Db 61 AWLEAQEE-BEVEGFPVTPQVPLRPMTYKAADVLSHFLKEKGGLEGLIHSQRQDILDWY 120  
QY 115 YHTQGYFPDQWNYTPGCVRYPLTFGWCYKLVPEPDKVEEANKGENTSLHVPVSLHGMD 174  
Db 121 YHTQGYFPDQWNYTPGCVRYPLTFGWCYKLVPEPDKVEEANKGENTSLHVPVSLHGMD 180  
QY 175 DPEREVLWRDLSRLAFHHVARELHPEYFKNC 206  
Db 181 DTEREVLWRDLSRLAFHHVARELHPEYFKNC 212

RESULT 18  
QOLJND  
nef protein - human immunodeficiency virus type 1 (isolate NDK)  
N:Alternate names: 3'-orf protein; orf-F protein  
C:Species: human immunodeficiency virus type 1, HIV-1  
A:Note: host Homo sapiens (man)  
C:Date: 30-Sep-1991 #sequence\_revision 30-Sep-1991 #text\_change 16-Jul-1999  
C:Accession: JQ0068  
R:Spire, B.; Sire, J.; Zachar, V.; Rey, F.; Barre-Sinoussi, F.; Galibert, F.; Hampe,  
Gene 81, 275-284, 1989  
A:Title: Nucleotide sequence of HIV1-NDK: a highly cytopathic strain of the human in  
A:Reference number: JQ0065; MUID:90034200  
A:Accession: JQ0068  
A:Molecule type: DNA  
A:Residues: 1-207 <SPI>  
A:Cross-references: GB:M27323; NID:g328154; PIDN:AAA44874.1; PID:g328163  
C:Genetics:  
A:Gene: nef  
C:Superfamily: AIDS nef protein  
C:Keywords: AIDS; immunodeficiency

Query Match 53.5%; Score 903.5; DB 1; Length 207;  
Best Local Similarity 77.3%; Pred. No. 1.1e-67;  
Matches 160; Conservative 25; Mismatches 21; Indels 1; Gaps 1;  
QY 1 MGGKWSKSSVVGWPTVRRMRRAEPAADGVGAASRDLEKHGAISSNTAATNAACAWLEA 60  
Db 1 MGGKWSKSLVGVKPAIRERIRKTDPAADGVGAASRDLEKHGAISSNTASTNDTCWLEA 60  
QY 61 QEE-BEVEGFPVTPQVPLRPMTYKAADVLSHFLKEKGGLEGLIHSQRQDILDWYHTQ 119  
Db 61 QEE-BEVEGFPVTPQVPLRPMTYKAADVLSHFLKEKGGLEGLIHSQRQDILDWYHTQ 120  
QY 120 YPDQWNYTPGCVRYPLTFGWCYKLVPEPDKVEEANKGENTSLHVPVSLHGMDPERE 179  
Db 121 YPDQWNYTPGCVRYPLTFGWCYKLVPEPDKVEEANKGENTSLHVPVSLHGMDPERE 180  
QY 180 VLEWRDLSRLAFHHVARELHPEYFKNC 206  
Db 181 VLMWRDLSRLAFHHVARELHPEYFKNC 207

RESULT 19  
B44963  
nef protein - human immunodeficiency virus type 1 (isolate Z321)  
N:Alternate names: 3'-orf protein; orf-F protein  
C:Species: human immunodeficiency virus type 1, HIV-1  
A:Note: host Homo sapiens (man)  
C:Date: 17-Feb-1994 #sequence\_revision 17-Feb-1994 #text\_change 28-May-1999

C:Accession: B44963  
R:Srinivasan, A.; York, D.; Butler Jr., D.; Jannoun-Nasr, R.; Getchell, J.; McCormick, J.  
AIDS Res. Hum. Retroviruses 5, 121-129, 1989  
A:Title: Molecular characterization of HIV-1 isolated from a serum collected in 1976: nu  
A:Reference number: A44963; MUID:89228766  
A:Accession: B44963  
A:Molecule type: DNA  
A:Residues: 1-205 <SR>  
A:Cross-references: GB:M15896; NID:g329392; PIDN:AAB53951.1; PID:g329397  
C:Genetics:  
A:Gene: nef; 3'-orf; orf-F  
C:Superfamily: AIDS nef protein  
C:Keywords: AIDS; immunodeficiency

Query Match 49.6%; Score 837; DB 1; Length 205;  
Best Local Similarity 73.2%; Pred. No. 3.8e-62;  
Matches 153; Conservative 23; Mismatches 25; Indels 8; Gaps 3;

QY 1 MGKWSKSSVVGWPTVRRMR---AEPADGVGAASRDLEKHGAISSNTAATNAACAW 57

Db 1 MGKWSKSSVVGWPTVRRMR---AEPADGVGAASRDLEKHGAISSNTAATNAACAW 56

QY 58 LEAQEE-EVGFPTVQVPLRPMYKAAVDLSHFLKEKGLGSLHSHQRQDILDWYH 116

Db 57 LEAQEE-EVGFPTVQVPLRPMYKAAVDLSHFLKEKGLGSLHSHQRQDILDWYH 116

QY 117 TQGYFDPWQNYTPGCVRYPLTFGCWYKLVPEPKVEANKGENTSLHPVSLHGMDDP 176

Db 117 TQGYFDPWQNYTPGCVRYPLTFGCWYKLVPEPKVEANKGENTSLHPVSLHGMDDP 176

QY 177 EREVLEWRFDLSRLAFHHVARELHPEYFKN 205

Db 177 EREVLEWRFDLSRLAFHHVARELHPEYFKN 205

## RESULT 20

T01673  
nef protein - human immunodeficiency virus type 1  
C:Species: human immunodeficiency virus type 1, HIV-1  
C:Date: 19-Feb-1999 #sequence\_revision 19-Feb-1999 #text\_change 21-Jul-2000  
C:Accession: T01673  
R:Alizon, M.; Wain-Hobson, S.; Gluckman, J.C.; Sonigo, P.  
Cell 46, 63-74, 1986  
A:Title: Genetic variability of the AIDS virus: Nucleotide sequence analysis of two isol  
A:Reference number: Z14389; MUID:86245056  
A:Accession: T01673  
A>Status: preliminary; translated from GB/EMBL/DBJ  
A:Molecule type: mRNA  
A:Residues: 1-209 <ALI>  
A:Cross-references: EMBL:K03456; NID:g60228; PIDN:CAA28017.1; PID:g60235  
C:Superfamily: AIDS nef protein

Query Match 49.4%; Score 834.5; DB 2; Length 209;  
Best Local Similarity 70.1%; Pred. No. 6.2e-62;  
Matches 148; Conservative 26; Mismatches 30; Indels 7; Gaps 2;

QY 1 MGKWSKSSVVGWPTVRRMRRAEPADGVG-----AASRDLEKHGAISSNTAATNAAC 55

Db 1 MGKWSKSSVVGWPTVRRMRRAEPADGVG-----AASRDLEKHGAISSNTAATNAAC 56

QY 56 AWEAQEE-EVGFPTVQVPLRPMYKAAVDLSHFLKEKGLGSLHSHQRQDILDWYH 115

Db 61 E--PPEEE-EVGFPTVQVPLRPMYKAAVDLSHFLKEKGLGSLHSHQRQDILDWYH 118

QY 116 HTQGYFDPWQNYTPGCVRYPLTFGCWYKLVPEPKVEANKGENTSLHPVSLHGMDD 175

Db 119 HTQGYFDPWQNYTPGCVRYPLTFGCWYKLVPEPKVEANKGENTSLHPVSLHGMDD 178

QY 176 PEREVLWRFDLSRLAFHHVARELHPEYFKN 206

Db 179 AEREVLKWKFDSSSLARHLHAREHQHPEYKDC 209

## RESULT 21

nef protein - simian immunodeficiency virus SIVcpz  
N:Alternate names: 3'-orf protein; orf-F protein  
C:Species: simian immunodeficiency virus SIVcpz  
A:Note: host Pan troglodytes (chimpanzee)  
C:Date: 30-Sep-1991 #sequence\_revision 30-Sep-1991 #text\_change 16-Jul-1999  
C:Accession: S09991  
R:Huet, T.; Cheynier, R.; Meyerhans, A.; Roelants, G.; Wain-Hobson, S.  
Nature 345, 356-359, 1990  
A:Title: Genetic organization of a chimpanzee lentivirus related to HIV-1.  
A:Reference number: S09983; MUID:90259077  
A:Accession: S09991  
A>Status: nucleic acid sequence not shown; translation not shown  
A:Molecule type: DNA  
A:Residues: 1-205 <HUE>  
A:Cross-references: EMBL:X52154; NID:g58866; PIDN:CAA36408.1; PID:g58877  
C:Genetics:  
A:Gene: nef  
C:Superfamily: AIDS nef protein  
C:Keywords: AIDS; immunodeficiency

Query Match 49.3%; Score 832; DB 1; Length 205;  
Best Local Similarity 70.2%; Pred. No. 9.8e-62;  
Matches 144; Conservative 27; Mismatches 34; Indels 0; Gaps 0;

QY 1 MGKWSKSSVVGWPTVRRMRRAEPADGVGAASRDLEKHGAISSNTAATNAACAWLEA 60

Db 1 MGTKWSKSSLVGWPEVRRIRREAPTAAGVGEVSKDLERHGAITSRTPTQTALWLEE 60

QY 61 QEEVEVGFPTVQVPLRPMYKAAVDLSHFLKEKGLGSLHSHQRQDILDWYHQTGY 120

Db 61 MDNEEVGFPTVQVPLRPMYKAAVDLSHFLKEKGLGSLHSHQRQDILDWYHQTGY 120

QY 121 FPDWQNYTPGCVRYPLTFGCWYKLVPEPKVEANKGENTSLHPVSLHGMDDPEREV 180

Db 121 FPDWQNYTPGCVRYPLTFGCWYKLVPEPKVEANKGENTSLHPVSLHGMDDPEREV 180

QY 181 LEWRFDLSRLAFHHVARELHPEYFKN 205

Db 181 LEWRFDLSRLAFHHVARELHPEYFKN 205

## RESULT 22

S33982  
trans-activating transcription regulator - human immunodeficiency virus type 1  
C:Species: human immunodeficiency virus type 1, HIV-1  
C:Date: 06-Oct-1994 #sequence\_revision 23-Feb-1996 #text\_change 20-Sep-1999  
C:Accession: S33982; S26385; S19864  
R:Carlini, F.  
submitted to the EMBL Data Library, November 1991  
A:Reference number: S33979  
A:Accession: S33982  
A:Molecule type: DNA  
A:Residues: 1-86 <CAR>  
A:Cross-references: EMBL:Z11530; NID:g60192; PIDN:CAA77625.1; PID:g60196  
R:Siderovski, D.P.; Matsuyama, T.; Frigerio, E.; Chui, S.; Min, X.; Erfe, H.; Summer  
Nucleic Acids Res. 20, 5311-5320, 1992  
A:Title: Random mutagenesis of the human immunodeficiency virus type-1 trans-activato  
A:Reference number: S26385; MUID:93065196  
A:Accession: S26385  
A:Molecule type: nucleic acid  
A:Residues: 1-86 <SID>  
A:Cross-references: EMBL:X64650; NID:g60144; PIDN:CAA45921.1; PID:g60145  
C:Genetics:  
A:Gene: tat  
A:Introns: 72/2  
C:Superfamily: AIDS trans-activating transcription regulator  
C:Keywords: AIDS; immunodeficiency

Query Match 28.8%; Score 486; DB 2; Length 86;  
Best Local Similarity 100.0%; Pred. No. 2e-33;  
Matches 85; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 209 EPVDPRLPKHPGSPKTAICTNCKKCCFHCQVCFITKALGISYGRKKRRRRPPQG 268  
|||||

Db 2 EPVDPRLPKHPGSPKTAICTNCKKCCFHCQVCFITKALGISYGRKKRRRRPPQG 61  
|||||

QY 269 SOTHQVSLSKQPTSQSRGDPGPK 293  
|||||

Db 62 SOTHQVSLSKQPTSQSRGDPGPK 86  
|||||

## RESULT 23

TNLJ12

trans-activating transcription regulator - human immunodeficiency virus type 1 (isolate  
C:Species: human immunodeficiency virus type 1, HIV-1

A:Note: host Homo sapiens (man)

C:Date: 04-Dec-1986 #sequence\_revision 04-Dec-1986 #text\_change 02-Jul-1998

C:Accession: A04017

R:Araya, S.K.; Gallo, R.C.

Proc. Natl. Acad. Sci. U.S.A. 83, 2209-2213, 1986

A:Title: Three novel genes of human T-lymphotropic virus type III: immune reactivity of  
A:Reference number: A94093; MUID:86177573

A:Accession: A04017

A:Molecule type: DNA

A:Residues: 1-95 <ARY>

C:Genetics:

A:Gene: tat

C:Superfamily: AIDS trans-activating transcription regulator

C:Keywords: AIDS: immunodeficiency; transcription regulation

Query Match 28.3%; Score 478; DB 1; Length 95;

Best Local Similarity 98.8%; Pred. No. 1e-32;

Matches 84; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 209 EPVDPRLPKHPGSPKTAICTNCKKCCFHCQVCFITKALGISYGRKKRRRRPPQG 268  
|||||

Db 11 EPVDPRLPKHPGSPKTAICTNCKKCCFHCQVCFITKALGISYGRKKRRRRPPQG 70  
|||||

QY 269 SOTHQVSLSKQPTSQSRGDPGPK 293  
|||||

Db 71 SOTHQVSLSKQPTSQSRGDPGPK 95  
|||||

## RESULT 24

A25700

trans-activating transcription regulator - human immunodeficiency virus type 1 (isolate

C:Species: human immunodeficiency virus type 1, HIV-1

C:Date: 28-Sep-1987 #sequence\_revision 28-Sep-1987 #text\_change 02-Jul-1998

C:Accession: A25700

R:Sodroski, J.; Patarca, R.; Rosen, C.; Wong-Staal, F.; Haseltine, W.

Science 229, 74-77, 1985

A:Reference number: A25700; MUID:85244627

A:Accession: A25700

A:Status: preliminary

A:Molecule type: mRNA

A:Residues: 1-86 <SOD>

C:Superfamily: AIDS trans-activating transcription regulator

Query Match 27.8%; Score 470; DB 2; Length 86;

Best Local Similarity 96.5%; Pred. No. 4.2e-32;

Matches 82; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

QY 209 EPVDPRLPKHPGSPKTAICTNCKKCCFHCQVCFITKALGISYGRKKRRRRPPQG 268  
|||||

Db 2 EPVDPRLPKHPGSPKTAICTNCKKCCFHCQVCFITKALGISYGRKKRRRRPPQG 61  
|||||

QY 269 SOTHQVSLSKQPTSQSRGDPGPK 293  
|||||

Db 62 SOTHQVSLSKQPTSQSRGDPGPK 86

## RESULT 25

S46353

nef protein - simian immunodeficiency virus SIVagm (isolate SAB-1)

C:Species: simian immunodeficiency virus SIVagm

A:Variety: isolate SAB-1

C:Date: 25-Dec-1994 #sequence\_revision 14-Feb-1997 #text\_change 20-Sep-1999

C:Accession: S46353

R:Jin, M.J.; Hui, H.; Robertson, D.L.; Mueller, M.C.; Barre-Sinoussi, F.; Hirsch, V.

EMBO J. 13, 2935-2947, 1994

A:Title: Mosaic genome structure of simian immunodeficiency virus from West African  
A:Reference number: S46335; MUID:94298785

A:Accession: S46353

A:Status: nucleic acid sequence not shown; translation not shown

A:Molecule type: DNA

A:Residues: 1-226 <JIN>

A:Cross-references: EMBL:U04005; NID:9466229; PIDN:AAA21511.1; PID:9466237

A:Experimental source: isolate SAB-1; sabaesus monkey

A:Note: the nucleotide sequence was submitted to the EMBL Data Library, December 1995

C:Genetics:

A:Gene: nef

C:Superfamily: AIDS nef protein

Query Match

Best Local Similarity 26.2%; Score 443; DB 2; Length 226;

Matches 98; Conservative 30; Mismatches 70; Indels 24; Gaps 6;

QY 1 MGKWSKS-----SVVGWPTVRR-MRRAPADGGAASRDLE-----KHGATSSNT 48  
|||||

Db 1 MGKSSKQQQRHSLWLWKLQAPVIQYDMLADPLGLQSGSHIQEACAKSLRDGLIRQGS 60  
|||||

QY 49 AATNAAC-----AWLEAOEEVEVFPVTPQVPLRPMTYKAAVDLSHFLKEKGLG 99  
|||||

Db 61 SRTSEGVKKHOGQPSWYD-EDEEVGFPVPCPLPRANTYKLAIDFGHFLKEKGLG 119  
|||||

QY 100 LIHSQRRQDILDIWIYHTQYFPDMQNYTPGCVRYPLTFGMCYKLVVPVDPKVEANKG 159  
|||||

Db 120 IYSESRKKILDLYALNENGVDMQNYTDGPGTRYPKCGMCKLVPV--DLSEEA 177  
|||||

QY 160 ENTSLHLPVSLHGMDDPEREVLWRFDSRLAFHHVARELHPE 201  
|||||

Db 178 ENHCLLHPAQVAYEDDANKETLVKFDPLAVDYVAVRLHPE 219  
|||||

## RESULT 26

S54385

nef protein - human immunodeficiency virus type 1 (fragment)

C:Species: human immunodeficiency virus type 1, HIV-1

C:Date: 15-Jul-1995 #sequence\_revision 01-Sep-1995 #text\_change 20-Sep-1999

C:Accession: S54385

R:Theodore, T.; Buckler-White, A.J.

submitted to the EMBL Data Library, July 1989

A:Reference number: S54377

A:Accession: S54385

A:Status: preliminary

A:Molecule type: genomic RNA

A:Residues: 1-97 <THE>

A:Cross-references: EMBL:M22639; NID:g329377; PIDN:AAA5371.1; PID:g555045

C:Superfamily: AIDS nef protein

Query Match

Best Local Similarity 26.0%; Score 439.5; DB 2; Length 97;

Matches 82; Conservative 6; Mismatches 8; Indels 1; Gaps 1;

QY 1 MGKWSKSSVVGWPTVRRMRRAEPADGGAASRDLEKHGATSSNTAATNACAWLEA 60  
|||||

Db 1 MGGRWSKSSIVGVPARIRIRTRDPAADGVGAVSRDLEKHGATSSNTRGTNADCAWLEA 60  
|||||

QY 61 QEE-EVGFPTVQVPLRPMTYKAAVDLSHFLKEKGG 96  
|||||







GenCore version 4.5  
Copyright (c) 1993 - 2000 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: August 26, 2002, 08:09:56 ; Search time 40.39 Seconds  
(without alignments)  
289.510 Million cell updates/sec

Title: US-09-509-239-13

Perfect score: 1688

Sequence: 1 MGCKSKSSVGVGPTVREM.....QSRGDPGPKETSGHHHHH 302

Scoring table:

Gapop 10.0 , Gapext 0.5

Searched: 105224 seqs, 38719550 residues

Total number of hits satisfying chosen parameters: 105224

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : SwissProt\_40:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

#### SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	1130	66.9	206	1	NEF_HV1BR P03406 human immun
2	1114	66.0	206	1	NEF_HV1PV P03405 human immun
3	1111	65.8	206	1	NEF_HV1W Q70627 human immun
4	1108	65.6	206	1	NEF_HV12 P03324 human immun
5	1088.5	64.5	205	1	NEF_HV1B8 P03855 human immun
6	1019	60.4	218	1	NEF_HV1B P12479 human immun
7	1005.5	59.6	205	1	NEF_HV1S3 P19545 human immun
8	997.5	59.1	211	1	NEF_HV1OY P20886 human immun
9	997	59.1	210	1	NEF_HV1A2 P03407 human immun
10	994	58.9	208	1	NEF_HV1RH P03858 human immun
11	991.5	58.7	214	1	NEF_HV1Y2 P35959 human immun
12	981	58.1	216	1	NEF_HV1JR P20867 human immun
13	974	57.7	208	1	NEF_HV1S1 P19546 human immun
14	927.5	54.9	206	1	NEF_HV1EL P04604 human immun
15	911	54.0	212	1	NEF_HV1Z6 P04602 human immun
16	903.5	53.5	207	1	NEF_HV1ND P18801 human immun
17	872	51.7	205	1	NEF_HV1U4 P24741 human immun
18	842	49.9	182	1	NEF_HV1MN P05856 human immun
19	837	49.6	205	1	NEF_HV1ZH P05859 human immun
20	834.5	49.4	209	1	NEF_HV1MA P04603 human immun
21	832	49.3	205	1	NEF_SVCZ P17664 chimpanzee
22	819	48.5	239	1	NEF_HV1C2 P05857 human immun
23	845	38.2	123	1	NEF_HV1H2 P04601 human immun
24	638	37.8	123	1	TAT_HV1B1 P03404 human immun
25	486	28.8	86	1	TAT_HV1B1 P04606 human immun
26	481	28.5	86	1	TAT_HV1PV P04607 human immun
27	478	28.3	86	1	TAT_HV112 P04326 human immun
28	470	27.8	86	1	TAT_HV1BR P04610 human immun
29	461	27.3	102	1	TAT_HV1RH P05908 human immun
30	454	26.9	86	1	TAT_HV1H2 P04608 human immun
31	446	26.4	101	1	TAT_HV1JR P20879 human immun
32	439.5	26.0	97	1	NEF_HV1Z2 P12478 human immun
33	437	25.9	101	1	TAT_HV1SC P05906 human immun

#### ALIGNMENTS

RESULT 1

NEF\_HV1BR

ID NEF\_HV1BR STANDARD; PRT: 206 AA.

AC P03406;

DT 21-JUL-1986 (Rel. 01, Created)

DT 21-JUL-1986 (Rel. 01, Last sequence update)

DT 16-OCT-2001 (Rel. 40, Last annotation update)

DE Negative factor (F-protein) (27 kDa protein) (3'ORF).

GN NEF.

OS Human immunodeficiency virus type 1 (BRU isolate) (HIV-1), and

OS Human immunodeficiency virus type 1 (New York-5 isolate) (HIV-1).

OC Viruses; Retroviral viruses; Retroviridae; Lentivirus.

OX NCBI\_TaxID=11686, 11698;

RN [1]

RP SEQUENCE FROM N.A.

RC STRAIN=ISOLATE BRU;

RX MEDLINE=8509333; PubMed=2981635;

RA Wain-Hobson S., Sonigo P., Danos O., Cole S., Alizon M.;

RT "Nucleotide sequence of the AIDS virus, LAV.;"

RL Cell 40:9-17(1985).

RN [2]

RP SEQUENCE FROM N.A. (CLONE PNL4-3).

RC STRAIN=ISOLATE NEW YORK-5;

RA Buckler C.E., Buckler-White A.J., Willey R.L., McCoy J.;

RL Submitted (JUN-1988) to the EMBL/GenBank/DBJ databases.

RN [3]

RP POST-TRANSLATIONAL MODIFICATIONS, AND FUNCTION.

RX MEDLINE=88039140; PubMed=3118220;

RA Guy B., Kieny M.-P., Riviere Y., le Peuch C., Dotti K., Girard M.;

RT "HIV F/3' orf encodes a phosphorylated GTP-binding protein resembling

an oncogene product.;"

RL Nature 330:286-289(1987).

RN [4]

RP X-RAY CRYSTALLOGRAPHY (2.5 ANGSTROMS) OF 54-205 IN COMPLEX WITH NEF.

RX MEDLINE=96279837; PubMed=8681387;

RA Lee C.H., Saksela K., Mirza U.A., Chait B.T., Kuriyan J.;

RT "Crystal structure of the conserved core of HIV-1 Nef complexed with

a Src family SH3 domain.;"

RL Cell 85:931-942(1996).

CC -!- FUNCTION: NEF HAS GTPASE, GTP-BINDING AND AUTOPHOSPHORYLATING

ACTIVITIES, IT SEEMS TO DOWN-REGULATE THE CD4(T4) ANTIGEN.

CC

CC This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL Outstation - the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See <http://www.isb-sib.ch/announce/> or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).

CC

CC EMBL; K02013; AAB59752.1; -

DR EMBL; M19921; AAA44993.1; -

DR EMBL; A04321; CAA00353.1; -

DR PIR; A04008; ASLJFV

DR PDB; IEFN; 1J-JAN-97.

DR HIV; K02013; NEFSBRU.  
DR HIV; M19921; NEFSNL43.  
DR InterPro: IPR001598; F-protein.  
DR Pfam: PF00469; F-protein; 1.  
KW AIDS; Myristate; GTP-binding; Phosphorylation; 3D-structure.  
FT LIPID 2 MYRISTATE.  
FT MOD\_RES 15 15 PHOSPHORYLATION (BY PKC).  
FT VARIANT 11 11 V -> I (IN CLONE PNL4-3).  
FT VARIANT 15 15 T -> A (IN CLONE PNL4-3).  
FT VARIANT 33 33 A -> V (IN CLONE PNL4-3).  
FT VARIANT 51 51 T -> N (IN CLONE PNL4-3).  
SQ SEQUENCE 206 AA; 23342 MW; 77453FC80B6004F2 CRC64;

Query Match 66.9%; Score 1130; DB 1; Length 206;  
Best Local Similarity 100.0%; Pred. No. 9.3e-92;  
Matches 206; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MGGKWSKSSVVGWPTVRMRRAEPAADGVGAASRDLEKKGATSSNTAATNAACAWLEA 60  
DB 1 MGGKWSKSSVVGWPTVRMRRAEPAADGVGAASRDLEKKGATSSNTAATNAACAWLEA 60  
QY 61 QEEVEGFPVTPQVPLRPMTYKAAVDLSHFLKEKGGLEGLIHSORRODILDLWIYHTQGY 120  
DB 61 QEEVEGFPVTPQVPLRPMTYKAAVDLSHFLKEKGGLEGLIHSORRODILDLWIYHTQGY 120  
QY 121 FPDQNTYTPGPGVRYPLTFGCYKLVPEPKVVEEANKGENTSLHPVSLHGMDDPREV 180  
DB 121 FPDQNTYTPGPGVRYPLTFGCYKLVPEPKVVEEANKGENTSLHPVSLHGMDDPREV 180  
QY 181 LEWRFDSRLAFHHVARELHPEYFKNC 206  
DB 181 LEWRFDSRLAFHHVARELHPEYFKNC 206

## RESULT 2

NEF\_HVIPV STANDARD; PRT; 206 AA.  
AC P03405;  
DT 21-JUL-1986 (Rel. 01, Created)  
DT 21-JUL-1986 (Rel. 01, Last sequence update)  
DT 16-OCT-2001 (Rel. 40, Last annotation update)  
DE Negative factor (F-protein) (27 kDa protein) (3'ORF).  
GN NEF.  
OS Human immunodeficiency virus type 1 (PV22 isolate) (HIV-1).  
OC Viruses; Retroviral viruses; Retroviridae; Lentivirus.  
OX NCBI\_TaxID=11700;  
RN [1]  
RP SEQUENCE FROM N.A.  
RX MEDLINE=85111157; Pubmed=2982104;  
RA Muesing M.A., Smith D.H., Cabradilla C.D., Benton C.V., Lasky L.A.,  
FA Capon D.J.;  
RT Nucleic acid structure and expression of the human  
RT AIDS/lymphadenopathy retrovirus.";  
RL Nature 313:450-458(1985).  
RN [2]  
RP POST-TRANSLATIONAL MODIFICATIONS, AND FUNCTION.  
RX MEDLINE=88039140; Pubmed=3118220;  
RA Guy B., Kieny M.-P., Riviere Y., le Peuch C., Dott K., Girard M.,  
RA Montagnier L., Lecocq J.-P.;  
RT "HIV F/3" orf encodes a phosphorylated GTP-binding protein resembling  
RT an oncogene product.";  
RL Nature 330:266-269(1987).  
CC -1- FUNCTION: NEF HAS GTPASE, GTP-BINDING AND AUTOPHOSPHORYLATING  
CC ACTIVITIES, IT SEEMS TO DOWN-REGULATE THE CD4(T4) ANTIGEN.  
CC  
CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
CC the European Bioinformatics Institute. There are no restrictions on its  
CC use by non-profit institutions as long as its content is in no way  
CC modified and this statement is not removed. Usage by and for commercial  
CC entities requires a license agreement (See <http://www.isb-sib.ch/announcement/>  
CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).

CC  
DR EMBL; K02083; AAB59874.1; -  
DR EMBL; X01762; -; NOT\_ANNOTATED\_CDS.  
DR PIR; A04007; ASLJVL.  
DR HSSP; P03406; IEFN.  
DR HIV; K02083; NEFSBV22.  
DR InterPro: IPR001558; F-protein.  
DR Pfam: PF00469; F-protein; 1.  
KW AIDS; Myristate; GTP-binding.  
FT LIPID 2 MYRISTATE.  
SQ SEQUENCE 206 AA; 23352 MW; EDE64281A17C6735 CRC64;

Query Match 66.0%; Score 1114; DB 1; Length 206;  
Best Local Similarity 97.6%; Pred. No. 2.3e-90;  
Matches 201; Conservative 3; Mismatches 2; Indels 0; Gaps 0;

QY 1 MGGKWSKSSVVGWPTVRMRRAEPAADGVGAASRDLEKKGATSSNTAATNAACAWLEA 60  
DB 1 MGGKWSKSSVVGWPTVRMRRAEPAADGVGAASRDLEKKGATSSNTAATNAACAWLEA 60  
QY 61 QEEVEGFPVTPQVPLRPMTYKAAVDLSHFLKEKGGLEGLIHSORRODILDLWIYHTQGY 120  
DB 61 QEEVEGFPVTPQVPLRPMTYKAAVDLSHFLKEKGGLEGLIHSORRODILDLWIYHTQGY 120  
QY 121 FPDQNTYTPGPGVRYPLTFGCYKLVPEPKVVEEANKGENTSLHPVSLHGMDDPREV 180  
DB 121 FPDQNTYTPGPGVRYPLTFGCYKLVPEPKVVEEANKGENTSLHPVSLHGMDDPREV 180  
QY 181 LEWRFDSRLAFHHVARELHPEYFKNC 206  
DB 181 LEWRFDSRLAFHHVARELHPEYFKNC 206

## RESULT 3

NEF\_HVILW STANDARD; PRT; 206 AA.  
AC Q70627;  
DT 15-JUL-1998 (Rel. 36, Created)  
DT 15-JUL-1998 (Rel. 36, Last sequence update)  
DT 16-OCT-2001 (Rel. 40, Last annotation update)  
DE Negative factor (F-protein) (27 kDa protein) (3'ORF).  
GN NEF.  
OS Human immunodeficiency virus type 1 (LW12.3 isolate) (HIV-1).  
OC Viruses; Retroviral viruses; Retroviridae; Lentivirus.  
OX NCBI\_TaxID=82834;  
RN [1]  
RP SEQUENCE FROM N.A.  
RX MEDLINE=95127297; Pubmed=7826699;  
RA Reitz M.S. Jr., Hall L., Robert-Guroff M., Lautenberger J., Hahn B.M.,  
RA Shaw G.M., Kong L.I., Weiss S.H., Waters D., Gallo R.C., Blattner W.;  
RT "Viral variability and serum antibody response in a laboratory worker  
RT infected with HIV type 1 (HTLV type IIIB).";  
RL AIDS Res. Hum. Retroviruses 10:1143-1155(1994).  
RN [2]  
RP STRUCTURE BY NMR OF 56-206.  
RX MEDLINE=97337445; Pubmed=9194185;  
RA Grzesiek S., Bax A., Hu J.S., Kaufman J., Palmer I., Stahl S.J.,  
RA Tjandra N., Wingfield P.T.;  
RT "Refined solution structure and backbone dynamics of HIV-1 Nef.";  
RL Protein Sci. 6:1248-1263(1997).  
CC -1- FUNCTION: NEF HAS GTPASE, GTP-BINDING AND AUTOPHOSPHORYLATING  
CC ACTIVITIES, IT SEEMS TO DOWN-REGULATE THE CD4(T4) ANTIGEN.  
CC  
CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
CC the European Bioinformatics Institute. There are no restrictions on its  
CC use by non-profit institutions as long as its content is in no way  
CC modified and this statement is not removed. Usage by and for commercial  
CC entities requires a license agreement (See <http://www.isb-sib.ch/announcement/>  
CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).

EMBL; U12055; AAA76691.1; -

DR PDB: 2NEF; 07-JUL-97.  
 DR InterPro: IPR001558; F-protein.  
 DR Pfam: PF00469; F-protein; 1.  
 KW AIDS; Myristate; GTP-binding; 3D-structure.  
 FT LIPID 2 MYRISTATE.  
 SQ SEQUENCE 206 AA; 23414 MW; ED81F68F6B61278E CRC64;

Query Match 65.8%; Score 1111; DB 1; Length 206;  
 Best Local Similarity 97.1%; Pred. No. 4.3e-90;  
 Matches 200; Conservative 5; Mismatches 1; Indels 0; Gaps 0;

Qy 1 MGGKWSKSSVVGWPTVRRMRRAEPAADGVGAASRDLEKKGAGTSSNTAATNAACAWLEA 60  
 Db 1 MGGKWSKSSVVGWPTVRRMRRAEPAADGVGAASRDLEKKGAGTSSNTAATNAACAWLEA 60  
 Qy 61 QEEEGVGPVTPQVPLRPMTYKAAVDLSHFLKEKGGLEGLHSQRQDILDLIWIHTQGY 120  
 Db 61 QEEEGVGPVTPQVPLRPMTYKAAVDLSHFLKEKGGLEGLHSQRQDILDLIWIHTQGY 120  
 Qy 121 FPDWQNYTPGGVRYPLTFGWCYKLVPEPKVEANKGENTSLHPVSLHGMDDPEREV 180  
 Db 121 FPDWQNYTPGGVRYPLTFGWCYKLVPEPKVEANKGENTSLHPVSLHGMDDPEREV 180  
 Qy 181 LEWRFDSRLAFHHVARELHPEYFKNC 206  
 Db 181 LEWRFDSRLAFHHVARELHPEYFKNC 206

RESULT 4  
 NEF\_HV112 STANDARD; PRT; 206 AA.  
 ID NEF\_HV112  
 AC P04324;  
 DT 20-MAR-1987 (Rel. 04, Created)  
 DT 13-AUG-1987 (Rel. 05, Last sequence update)  
 DT 16-OCT-2001 (Rel. 40, Last annotation update)  
 DE Negative factor (F-protein) (27 kDa protein) (3'ORF).  
 GN NEF.  
 OS Human immunodeficiency virus type 1 (clone 12) (HIV-1).  
 OC Viruses; Retroviral viruses; Retroviridae; Lentivirus.  
 OX NCBI\_TaxId=11679;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RX MEDLINE=86177573; PubMed=3008154;  
 RA Arya S.K., Gallo R.C.;  
 RT "Three novel genes of human T-lymphotropic virus type III: Immune  
 RT reactivity of their products with sera from acquired immune  
 RT deficiency syndrome patients."  
 RL Proc. Natl. Acad. Sci. U.S.A. 83:2209-2213(1986).  
 RN [2]  
 RP POST-TRANSLATIONAL MODIFICATIONS, AND FUNCTION.  
 RX MEDLINE=88039140; PubMed=3118220;  
 RA Guy B., Kieny M.-P., Riviere Y., le Peuch C., Dott K., Girard M.,  
 RA Montagnier L., Lecocq J.-P.;  
 RT "HIV F/3' orf encodes a phosphorylated GTP-binding protein resembling  
 RT an oncogene product."  
 RL Nature 330:266-269(1987).  
 CC -!- FUNCTION: NEF HAS GTPASE, GTP-BINDING AND AUTOPHOSPHORYLATING  
 CC ACTIVITIES, IT SEEMS TO DOWN-REGULATE THE CD4(T4) ANTIGEN.  
 CC  
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation  
 CC the European Bioinformatics Institute. There are no restrictions on its  
 CC use by non-profit institutions as long as its content is in no way  
 CC modified and this statement is not removed. Usage by and for commercial  
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
 CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
 CC  
 CC EMBL; M11840; AAA45001.1;  
 DR PIR; A04006; ASLJ12.  
 DR HSSP; P03406; IEFN.  
 DR HIV; M11840; NEFSPC12.  
 DR InterPro; IPR001558; F-protein.

DR Pfam: PF00469; F-protein; 1.  
 KW AIDS; Myristate; GTP-binding.  
 FT LIPID 2 MYRISTATE.  
 SQ SEQUENCE 206 AA; 23366 MW; 218F5B2980F79A46 CRC64;

Query Match 65.6%; Score 1108; DB 1; Length 206;  
 Best Local Similarity 97.1%; Pred. No. 7.8e-90;  
 Matches 200; Conservative 4; Mismatches 2; Indels 0; Gaps 0;

Qy 1 MGGKWSKSSVVGWPTVRRMRRAEPAADGVGAASRDLEKKGAGTSSNTAATNAACAWLEA 60  
 Db 1 MGGKWSKSSVVGWPTVRRMRRAEPAADGVGAASRDLEKKGAGTSSNTAATNAACAWLEA 60  
 Qy 61 QEEEGVGPVTPQVPLRPMTYKAAVDLSHFLKEKGGLEGLHSQRQDILDLIWIHTQGY 120  
 Db 61 QEEEGVGPVTPQVPLRPMTYKAAVDLSHFLKEKGGLEGLHSQRQDILDLIWIHTQGY 120  
 Qy 121 FPDWQNYTPGGVRYPLTFGWCYKLVPEPKVEANKGENTSLHPVSLHGMDDPEREV 180  
 Db 121 FPDWQNYTPGGVRYPLTFGWCYKLVPEPKVEANKGENTSLHPVSLHGMDDPEREV 180  
 Qy 181 LEWRFDSRLAFHHVARELHPEYFKNC 206  
 Db 181 LEWRFDSRLAFHHVARELHPEYFKNC 206

RESULT 5  
 NEF\_HV1B8 STANDARD; PRT; 205 AA.  
 ID NEF\_HV1B8  
 AC P05855;  
 DT 01-NOV-1988 (Rel. 09, Created)  
 DT 01-NOV-1988 (Rel. 09, Last sequence update)  
 DT 16-OCT-2001 (Rel. 40, Last annotation update)  
 DE Negative factor (F-protein) (27 kDa protein) (3'ORF).  
 GN NEF.  
 OS Human immunodeficiency virus type 1 (BH8 isolate) (HIV-1).  
 OC Viruses; Retroviral viruses; Retroviridae; Lentivirus.  
 OX NCBI\_TaxId=11684;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RX MEDLINE=85111123; PubMed=2578615;  
 RA Ratner L., Haseltine W., Patarca R., Livak K.J., Starcich B.R.,  
 RA Josephs S.F., Doran E.R., Rafalski J.A., Whitehorn E.A.,  
 RA Baumanster K., Ivanoff L., Petteway S.R. Jr., Pearson M.L.,  
 RA Lautenberger J.A., Papas T.S., Grayeb J., Chang N.T., Gallo R.C.,  
 RA Wong-Staal F.;  
 RT "Complete nucleotide sequence of the AIDS virus, HTLV-III."  
 RL Nature 313:277-284(1985).  
 RN [2]  
 RP POST-TRANSLATIONAL MODIFICATIONS, AND FUNCTION.  
 RX MEDLINE=88039140; PubMed=3118220;  
 RA Guy B., Kieny M.-P., Riviere Y., le Peuch C., Dott K., Girard M.,  
 RA Montagnier L., Lecocq J.-P.;  
 RT "HIV F/3' orf encodes a phosphorylated GTP-binding protein resembling  
 RT an oncogene product."  
 RL Nature 330:266-269(1987).  
 CC -!- FUNCTION: NEF HAS GTPASE, GTP-BINDING AND AUTOPHOSPHORYLATING  
 CC ACTIVITIES, IT SEEMS TO DOWN-REGULATE THE CD4(T4) ANTIGEN.  
 CC  
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation  
 CC the European Bioinformatics Institute. There are no restrictions on its  
 CC use by non-profit institutions as long as its content is in no way  
 CC modified and this statement is not removed. Usage by and for commercial  
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
 CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
 CC  
 CC EMBL; K02011; -; NOT\_ANNOTATED\_CDS.  
 DR HSSP; P03406; IEFN.  
 DR HIV; K02011; NEF5B8.  
 DR InterPro; IPR001558; F-protein.  
 DR Pfam; PF00469; F-protein; 1.

```
KW AIDS: Myristate; GTP-binding.
FT LIPID 2 MYRISTATE.
SQ SEQUENCE 205 AA; 23305 MW; 8BC12F6650DD111 CRC64;

Query Match 64.5%; Score 1088.5; DB 1; Length 205;
Best Local Similarity 96.6%; Pred. No. 3.9e-88;
Matches 199; Conservative 3; Mismatches 3; Indels 1; Gaps 1;

Qy 1 MGGKWSKSSVVGWPTVTRMRRAEPAADGVGAASRDLEKKGAISSNTAATNAACAWLEA 60
Db 1 MGGKWSKSSVVGWPAVRMRRAEPAADGVGAASRDLEKKGAISSNTAATNAACAWLEA 60
Qy 61 QEEVEVGFPVTPQVPLRPMTYKAAVDLSHFLKEKGGLEGLIHSORRQDILDLWIHTQGY 120
Db 61 QEEVEVGFPVTPQVPLRPMTYKAAVDLSHFLKEKGGLEGLIHSORRQDILDLWIHTQGY 120
Qy 121 FPDQNYTPGQVRYPLTFGWCYKLVPEPKVEEANKGENTSLHPVSLHGMDDPEREV 180
Db 121 FPDQNYTPGQVRYPLTFGWCYKLVPEPEK-EEANKGENTSLHPVSLHGMDDPEREV 179
Qy 181 LEWRDLSRLAFHHVARELHPEYFKNC 206
Db 180 LEWRDLSRLAFHHMARELHPEYFKNC 205

RESULT 6
NEF_HV1BN STANDARD; PRT; 218 AA.
AC P12479;
DT 01-OCT-1989 (Rel. 12, Created)
DT 01-OCT-1989 (Rel. 12, Last sequence update)
DT 16-OCT-2001 (Rel. 40, Last annotation update)
DE Negative factor (F-protein) (27 kDa protein) (3'ORF).
GN NEF.
OS Human immunodeficiency virus type 1 (LBR isolate) (HIV-1).
OC Viruses; Retroviral viruses; Retroviridae; Lentivirus.
OX NCBI_TaxID=11693;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=89083613; PubMed=2789516;
RA Anand R., Thayer R., Srinivasan A., Nayyar S., Gardner M., Luciw P.,
RA Dandekar S.;
RT "Biological and molecular characterization of human immunodeficiency
RT virus (HIV-LBR) from the brain of a patient with progressive
RT dementia.";
RL Virology 168:79-89(1989).
CC -!- FUNCTION: NEF HAS GTPASE, GTP-BINDING AND AUTOPHOSPHORYLATING
CC ACTIVITIES, IT SEEMS TO DOWN-REGULATE THE CD4(T4) ANTIGEN.
CC -!- MISCELLANEOUS: THIS VIRUS IS CYTOPATHICALLY ACTIVE AND WAS
CC HARVESTED FROM THE BRAIN TISSUE OF A NEUROLOGICAL AIDS PATIENT.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announcement/
CC or send an email to license@isb-sib.ch).
CC -----
CC EMBL; M21098; AAA44222.1; -.
CC PIR; D31667; ASLJBR.
CC HSP; P03406; 1EPN.
CC HIV; M21098; NEFSBRVA.
CC InterPro; IPR001558; F-protein.
CC Pfam; PF00469; F-protein; 2.
KW AIDS: Myristate; GTP-binding.
FT LIPID 2 MYRISTATE (BY SIMILARITY).
SQ SEQUENCE 218 AA; 25032 MW; FCADAE1045C460E CRC64;

Query Match 60.4%; Score 1019; DB 1; Length 218;
Best Local Similarity 84.1%; Pred. No. 5e-82;
```

```
Matches 185; Conservative 12; Mismatches 7; Indels 16; Gaps 2;

Qy 1 MGGKWSKSSVVGWPTVTRMR-----RAEPAADGVGAASRDLEKKGAISS 46
Db 1 MGGKWSK--MAGWSTVTRMRRAEPAARERMRRAEPAADGVGAASRDLEKKGAISS 58
Qy 47 NTAATNAACAWLEAQEEVEVGFPVTPQVPLRPMTYKAAVDLSHFLKEKGGLEGLIHSORR 106
Db 59 NTAATNACAWLEAQEEVEVGFPVTPQVPLRPMTYKAAVDLSHFLKEKGGLEGLIHSORR 118
Qy 107 QDILDLWIHTQGYFPDQNYTPGQVRYPLTFGWCYKLVPEPKVEEANKGENTSLH 166
Db 119 QDILDLWIHTQGYFPDQNYTPGQVRYPLTFGWCYKLVPEPEKIEEANGENNSLH 178
Qy 167 PVSILHGMDDPEREVLEWRDLSRLAFHHVARELHPEYFKNC 206
Db 179 PMSQHGMDDPEREVLEWRDLSRLAFHHMARELHPEYFKNC 218

RESULT 7
NEF_HV1S3 STANDARD; PRT; 205 AA.
AC P19545;
DT 01-FEB-1991 (Rel. 17, Created)
DT 01-FEB-1991 (Rel. 17, Last sequence update)
DT 16-OCT-2001 (Rel. 40, Last annotation update)
DE Negative factor (F-protein) (27 kDa protein) (3'ORF).
GN NEF.
OS Human immunodeficiency virus type 1 (SF33 isolate) (HIV-1).
OC Viruses; Retroviral viruses; Retroviridae; Lentivirus.
OX NCBI_TaxID=11690;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=90317906; PubMed=2370688;
RA York-Higgins D., Cheng-Mayer C., Bauer D., Levy J.A., Dina D.;
RT "Human immunodeficiency virus type 1 cellular host range,
RT replication, and cytopathicity are linked to the envelope region of
RT the viral genome.";
RL J. Virol. 64:4016-4020(1990).
CC -!- FUNCTION: NEF HAS GTPASE, GTP-BINDING AND AUTOPHOSPHORYLATING
CC ACTIVITIES, IT SEEMS TO DOWN-REGULATE THE CD4(T4) ANTIGEN.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announcement/
CC or send an email to license@isb-sib.ch).
CC -----
CC EMBL; M38427; AAA45068.1; -.
CC HSP; P03406; 1EPN.
CC HIV; M38427; NEFSF33.
CC InterPro; IPR001558; F-protein.
CC Pfam; PF00469; F-protein; 1.
KW AIDS: Myristate; GTP-binding.
FT LIPID 2 MYRISTATE (BY SIMILARITY).
SQ SEQUENCE 205 AA; 23318 MW; FFB419A1C5DFC9F3 CRC64;

Query Match 59.6%; Score 1005.5; DB 1; Length 205;
Best Local Similarity 86.4%; Pred. No. 7e-81;
Matches 178; Conservative 19; Mismatches 8; Indels 1; Gaps 1;

Qy 1 MGGKWSKSSVVGWPTVTRMRRAEPAADGVGAASRDLEKKGAISSNTAATNAACAWLEA 60
Db 1 MGGKWSKSK-MGWPAVRMRRAEPAADGVGAASRDLEKKGAISSNTAATNAACAWLEA 59
Qy 61 QEEVEVGFPVTPQVPLRPMTYKAAVDLSHFLKEKGGLEGLIHSORRQDILDLWIHTQGY 120
Db 60 QEEVEVGFPVTPQVPLRPMTYKAAVDLSHFLKEKGGLEGLVSKRQDILDLWIHTQGY 119
Qy 121 FPDQNYTPGQVRYPLTFGWCYKLVPEPKVEEANKGENTSLHPVSLHGMDDPEREV 180
```



AC P05858;  
 DT 01-NOV-1988 (Rel. 09, Created)  
 DT 01-NOV-1988 (Rel. 09, Last sequence update)  
 DT 16-OCT-2001 (Rel. 40, Last annotation update)  
 DE Negative factor (F-protein) (27 kDa protein) (3'ORF).  
 GN NEF.  
 OS Human immunodeficiency virus type 1 (RE/HAT isolate) (HIV-1).  
 OC Viruses; Retrovirdae; Retroviridae; Lentivirus.  
 OX NCBI\_TaxID=11701;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RA Starich B.R., Hahn B.H., Shaw G.M., McNeely P.D., Modrow S.,  
 RA Wolf H., Parks E.S., Parks W.P., Josephs S.F., Gallo R.C.,  
 RA Wong-Staal F.;  
 RL Submitted (XXX-1987) to the HIV data bank.  
 CC -!- FUNCTION: NEF HAS GTPASE, GTP-BINDING AND AUTOPHOSPHORYLATING  
 CC ACTIVITIES, IT SEEMS TO DOWN-REGULATE THE CD4(T4) ANTIGEN.  
 CC  
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
 CC the European Bioinformatics Institute. There are no restrictions on its  
 CC use by non-profit institutions as long as its content is in no way  
 CC modified and this statement is not removed. Usage by and for commercial  
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announcement/>  
 CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
 CC  
 DR EMBL; M17451; AAA45058.1;  
 DR HSP; P03406; IEFN.  
 DR HIV; M17451; NEFSRF.  
 DR InterPro: IPR001558; F-protein.  
 DR Pfam: PF00469; F-protein; 1.  
 DR ADS; Myristate; GTP-binding.  
 KW ADS; Myristate; GTP-binding.  
 FT LIPID 2  
 SQ SEQUENCE 208 AA; 23532 MW; 8F836FE980F084C CRC64;  
 Query Match 58.9%; Score 994; DB 1; Length 208;  
 Best Local Similarity 84.6%; Pred. No. 7.3e-80;  
 Matches 176; Conservative 16; Mismatches 14; Indels 2; Gaps 1;  
 QY 1 MGGKWSKSSVVGWPTVR-----ERMRRAEPAADGVGAASRDLEKKGALTSNTAA 60  
 DB 1 MGGKWSKSKMGWPAVRERMRRAEPAADGVGAASRDLEKKGALTSNTAA 60  
 QY 61 Q--EEVEGFPVTPQVPLRPTMYTAAVDLSHFLKKGLEGLIHSQRQDILDWYHTQ 118  
 DB 61 QEEDEEYGFVTPQVPLRPTMYTAAVDLSHFLKKGLEGLIHSQRQDILDWYHTQ 120  
 QY 119 GFDPQWNTPGVRYPLTFTGWCYKLVPEPKVEEANKGENTSLHHPVSLHGMDPDER 178  
 DB 121 GFDPQWNTPGVRYPLTFTGWCYKLVPEPKVEEANKGENTSLHHPVSLHGMDPDER 180  
 QY 179 EVLEWRDRLAFHVAHRELHPEYKNC 206  
 DB 181 EVLVNFKDRLAFHVAHRELHPEYKNC 208  
 RESULT 11  
 NEF\_HV1Y2  
 ID NEF\_HV1Y2 STANDARD; PRT; 214 AA.  
 AC P35959;  
 DT 01-JUN-1994 (Rel. 29, Created)  
 DT 01-JUN-1994 (Rel. 29, Last sequence update)  
 DT 16-OCT-2001 (Rel. 40, Last annotation update)  
 DE Negative factor (F-protein) (27 kDa protein) (3'ORF).  
 GN NEF.  
 OS Human immunodeficiency virus type 1 (YU-2 isolate) (HIV-1).  
 OC Viruses; Retrovirdae; Retroviridae; Lentivirus.  
 OX NCBI\_TaxID=36377;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RA Li Y., Hui H., Burgess C.J., Price R.W., Sharp P.M., Hahn B.H.,

Shaw G.M.;  
 RT Complete nucleotide sequence, genome organization, and biological  
 RT properties of human immunodeficiency virus type 1 in vivo: evidence  
 RT for limited defectiveness and complementation.;  
 RL J. Virol. 66:6587-6600(1992).  
 CC -!- FUNCTION: NEF HAS GTPASE, GTP-BINDING AND AUTOPHOSPHORYLATING  
 CC ACTIVITIES, IT SEEMS TO DOWN-REGULATE THE CD4(T4) ANTIGEN.  
 CC  
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
 CC the European Bioinformatics Institute. There are no restrictions on its  
 CC use by non-profit institutions as long as its content is in no way  
 CC modified and this statement is not removed. Usage by and for commercial  
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announcement/>  
 CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
 CC  
 DR EMBL; M93258; -; NOT\_ANNOTATED\_CDS.  
 DR PIR; I44001; I44001.  
 DR HSP; P03406; IEFN.  
 DR InterPro: IPR001558; F-protein.  
 DR Pfam: PF00469; F-protein; 1.  
 DR ADS; Myristate; GTP-binding.  
 KW ADS; Myristate; GTP-binding.  
 FT LIPID 2  
 SQ SEQUENCE 214 AA; 24532 MW; E188D43D7B084D04 CRC64;  
 Query Match 58.7%; Score 991.5; DB 1; Length 214;  
 Best Local Similarity 84.2%; Pred. No. 1.2e-79;  
 Matches 181; Conservative 12; Mismatches 11; Indels 11; Gaps 2;  
 QY 1 MGGKWSKSSVVGWPTVR-----ERMRRAEPAADGVGAASRDLEKKGALTSNTAA 50  
 DB 1 MGGKWSKSKMGWPAVRERMRRAEPAADGVGAASRDLEKKGALTSNTAA 60  
 QY 51 TNAACAWLEAQEEEEVEGFPVTPQVPLRPTMYTAAVDLSHFLKKGLEGLIHSQRQDIL 110  
 DB 61 TNAACAWLEAQEEEEVEGFPVTPQVPLRPTMYTAAVDLSHFLKKGLEGLIHSQRQDIL 120  
 QY 111 DLWYHTQGYFPDQWNTPGVRYPLTFTGWCYKLVPEPKVEEANKGENTSLHHPVSL 170  
 DB 121 DLWYHTQGYFPDQWNTPGVRYPLTFTGWCYKLVPEPKVEEANKGENTSLHHPVSL 179  
 QY 171 HGMDPDERLEWRDRLAFHVAHRELHPEYKNC 205  
 DB 180 HGMDPDERLEWRDRLAFHVAHRELHPEYKNC 214  
 RESULT 12  
 NEF\_HV1JR  
 ID NEF\_HV1JR STANDARD; PRT; 216 AA.  
 AC P20867;  
 DT 01-FEB-1991 (Rel. 17, Created)  
 DT 01-FEB-1991 (Rel. 17, Last sequence update)  
 DT 16-OCT-2001 (Rel. 40, Last annotation update)  
 DE Negative factor (F-protein) (27 kDa protein) (3'ORF).  
 GN NEF.  
 OS Human immunodeficiency virus type 1 (JRCSE isolate) (HIV-1).  
 OC Viruses; Retrovirdae; Retroviridae; Lentivirus.  
 OX NCBI\_TaxID=11688;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RA Koyanagi S., Chen I.S.Y.;  
 RL Submitted (DEC-1988) to the HIV data bank.  
 CC -!- FUNCTION: NEF HAS GTPASE, GTP-BINDING AND AUTOPHOSPHORYLATING  
 CC ACTIVITIES, IT SEEMS TO DOWN-REGULATE THE CD4(T4) ANTIGEN.  
 CC  
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
 CC the European Bioinformatics Institute. There are no restrictions on its  
 CC use by non-profit institutions as long as its content is in no way  
 CC modified and this statement is not removed. Usage by and for commercial  
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announcement/>  
 CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).

```

CC
EMBL; M38429; AAB03750.1; -.
DR
HSSP; P03406; 1EFN.
DR
HIV; M38429; NEFSJRCSF.
DR
InterPro; IPR001558; F-protein.
DR
Pfam; PF00469; F-protein; 1.
DR
AIDS; Myristate; GTP-binding.
KW
LIPID 2
FT
SEQUENCE 216 AA; 24567 MW; D163FFA8C71529DC CRC64;
SQ
Query Match 58.1%; Score 981; DB 1; Length 216;
Best Local Similarity 81.9%; Pred. No. 1e-78;
Matches 177; Conservative 15; Mismatches 14; Indels 10; Gaps 1;
QY 1 MGKWKSSVYGVGPTVRRMRRAEPAAD-----GVGAASRDLEKHGAITSSNTAA 50
DB 1 MGKWKSKHSPGVGKSTVRRMRRAEPATDRVQTPEAAVGVGAVSRDLEKHGAITSSNTAA 60
QY 51 TNAACAWLEAQEEVEEFGFPTQVPLRPMTYKAAVDLSHLFLKEKGGLEGLHSORRQDIL 110
DB 61 TNADCAWLEAYEEDVEEFGFVPRPQVPLRPMTYKAAIDLSHLFLKEKGGLEGLYSOKRDIL 120
QY 111 DLWYHTQGVFPDQWNTYTPGQVRYPLTFGWCYKLVPEPKVDEANKGENTSLLHPVSL 170
DB 121 DLWYHTQGVFPDQWNTYTAGVRFPLTFGWCFLVLPVDPKVEAEENNCLLHPMSQ 180
QY 171 HGMDPPEVLEWFRDSRLAPHHVARELHPPEYFKNC 206
DB 181 HGMDPPEVLEWFRDSRLAPHHVARELHPPEYFKNC 216
RESULT 13
NEF_HV1S1 STANDARD; PRT; 208 AA.
ID ID NEF_HV1S1
AC AC P19546;
DT 01-FEB-1991 (Rel. 17, Created)
DT 01-FEB-1991 (Rel. 17, Last sequence update)
DT 16-OCT-2001 (Rel. 40, Last annotation update)
DE Negative factor (F-protein) (27 kDa protein) (3'ORF).
GN NEF.
OS Human immunodeficiency virus type 1 (SF162 isolate) (HIV-1).
OC Viruses; Retroid viruses; Retroviridae; Lentivirus.
OX NCBI_TaxID=11691;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=90347835; PubMed=2384920;
RA Cheng-Mayer C., Quiroga M., Tung J.W., Dina D., Levy J.;
RT "Viral determinants of human immunodeficiency virus type 1 T-cell or
RT macrophage tropism, cytopathogenicity, and CD4 antigen modulation.";
RL J. Virol. 64:4390-4398(1990).
CC CC -1- FUNCTION: NEF HAS GTPASE, GTP-BINDING AND AUTOPHOSPHORYLATING
CC ACTIVITIES, IT SEEMS TO DOWN-REGULATE THE CD4(T4) ANTIGEN.
CC
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announcement/);
CC or send an email to license@isb-sib.ch).
CC -----
CC EMBL; M65024; AAA45073.1; -.
CC HSSP; P03406; 1EFN.
CC HIV; M38428; NEFSF162.
CC InterPro; IPR001558; F-protein.
CC Pfam; PF00469; F-protein; 1.
CC AIDS; Myristate; GTP-binding.
KW
LIPID 2
FT
SEQUENCE 208 AA; 23684 MW; AOB1007D14E46E32 CRC64;
SQ
MYRISTATE (BY SIMILARITY).

```

Query Match 57.7%; Score 974; DB 1; Length 208;

```

Best Local Similarity   84.2%; Pred. No. 4.1e-78;
Matches 176; Conservative      Mismatches 13; Indels 4; Gaps 2;

Qy  1  MGKWKSSVVGNPTVRRMR---APPAADGVGAASRDLEKHGCAITSSNTAATAACAW 57
    ||||| : ||||| : ||||| : ||||| : ||||| : ||||| : ||||| : |||||
Db  1  MGKWSK-RWSGWSAVRERMKRAEPAADGVGAASRDLEKHGCAITSSNTAANNADCAW 59

Qy  58 LEAQEESEVGFPVPQVPLRPMTYKAADVLSHFKEKGGLGLELTHSORRODILWIYHT 117
    ||||| : ||||| : ||||| : ||||| : ||||| : ||||| : ||||| : |||||
Db  60 LEAQEDVDGFPVPQVPLRPMTYKAALDLSHFLKEKGGLGLELYSOKRKQDILDWIHT 119

Qy  118 QGYFDPQNQTTPGGVARYPLTFGCYKLVPVEPKVEANKGENTSLHPVSLHGMDDPE 177
    ||||| : ||||| : ||||| : ||||| : ||||| : ||||| : ||||| : |||||
Db  120 QGYFDPQNQTTPGGIRYPLTFGCWFKLVPPDYVEANAGENNSSLHPMSQHGMDDPE 179

Qy  178 REVLWFDSRLAFHHVARELHPEYFKNC 206
    :|| : ||||| : ||||| : ||||| : ||||| : ||||| : ||||| : |||||
Db  180 KEVLWRFDSRLAFHHMARELHPEYKDC 208

RESULT 14
NEF_HVIEL STANDARD; PRT; 206 AA.
ID ID NEF_HVIEL AC P04604;
DT DT 13-AUG-1987 (Rel. 05, Created)
DT DT 13-AUG-1987 (Rel. 05, Last sequence update)
DT DT 16-OCT-2001 (Rel. 40, Last annotation update)
DT DT Negative factor (F-protein) (27 kDa protein) (3'ORF).
CN NEF.
OS OS Human immunodeficiency virus type 1 (ELI isolate) (HIV-1).
OC OC Viruses; Retroid viruses; Retroviridae; Lentivirus.
ON NCBI_TaxID=11689;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=86245056; PubMed=2424612;
RA Alizon M., Wain-Hobson S., Montagnier L., Sonigo P.;
RT "Genetic variability of the AIDS virus: nucleotide sequence analysis
RT of two isolates from African patients.";
RL Cell 46:63-74(1986).
RN [2]
RP POST-TRANSLATIONAL MODIFICATIONS, AND FUNCTION.
RX MEDLINE=88039140; PubMed=3118220;
RA Guy B., Kieny M.-P., Riviere Y., le Peuch C., Dott K., Girard M.,
RA Montagnier L., Lecocq J.-P.;
RT "HIV F3'orf encodes a phosphorylated GTP-binding protein resembling
RT an oncogene product.";
RL Nature 330:266-269(1987).
CC CC -!- FUNCTION: NEF HAS GTPASE, GTP-BINDING AND AUTOPHOSPHORYLATING
CC CC ACTIVITIES, IT SEEMS TO DOWN-REGULATE THE CD4(T4) ANTIGEN.
-----
CC CC THIS SWISS-PROT entry is copyright. It is produced through a collaboration
CC CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC CC the European Bioinformatics Institute. There are no restrictions on its
CC CC use by non-profit institutions as long as its content is in no way
CC CC modified and this statement is not removed. Usage by and for commercial
CC CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC CC or send an email to license@isb-sib.ch).
-----
DR EMBL; K03454; AAA4330.1; -.
DR EMBL; A07108; CAA00617.1; -.
DR HSP; P03406; IEFN.
DR HIV; K03454; NEF$ELI.
DR InterPro; IPR001558; F-protein.
DR Pfam; PF00469; F-protein; 1.
KW AIDS; Myristate; GTP-binding.
FT LIPID 2 MYRISTATE.
SQ SEQUENCE 206 AA; 23612 MW; 4CFF9F18AEAB503C CRC64;
```

Query Match	54.9%	Score 927.5;	DB 1;	Length 206;
Best Local Similarity	79.1%	Pred. NO. 4.7e-74;		
Matches 163; Conservative	24;	Mismatches 18;	Indels 1;	Gaps 1;



QY 1 MGGKWSKSSVVGWPTVRRMRRAEPAADGVGAASRDLEKKGAISSNTAATNACAWLEA 60  
DB 1 MGGKWSKSSVVGWPAIRIRIRRTNPAADGVGAASRDLEKKGAISSNTASTNADCAWLEA 60  
QY 61 QEE-EEVGFPPVTPQVPLRPMTYKAAVDLSHFLKKGGLGLIHSORRODILWYHTQG 119  
DB 61 QEESEVGFPPVTPQVPLRPMTYKAAVDLSHFLKKGGLGLIHSORRODILWYHTQG 120  
QY 120 YFPDQNYTPGPGVRYPLTFGWCYKLVPEPKVEANKGENTSLHVPVSLHGMDPERE 179  
DB 121 IFPDQNYTPGPGIRYPLTFGWCYKLVPEPKVEANKGENTSLHVPVSLHGMDPERE 180  
QY 180 VLEWRDRLAFHVAARELHPEYKNC 205  
DB 181 VLKWRFRNSRLAFEHKAREMHPEFYKN 206

RESULT 15  
ID\_NEF\_HV126 STANDARD; PRT; 212 AA.  
AC P04602;  
DT 13-AUG-1987 (Rel. 05, Created)  
DT 13-AUG-1987 (Rel. 05, Last sequence update)  
DT 16-OCT-2001 (Rel. 40, Last annotation update)  
DE Negative factor (F-protein) (27 kDa protein) (3'ORF).  
GN NEF.  
OS Human immunodeficiency virus type 1 (Zaire 6 isolate) (HIV-1).  
OC Viruses; Retroviral viruses; Retroviridae; Lentivirus.  
OX NCBI\_TaxID=11708;  
RN [1]  
RP SEQUENCE FROM N.A.  
RX MEDLINE-87248097; PubMed=3036660;  
RA Srinivasan A., Anand R., York D., Ranganathan P., Feorino P.,  
RA Schochetman G., Curran J., Kalyanaram V.S., Luciw P.A.,  
RA Sanchez-Pescador R.;  
RT "Molecular characterization of human immunodeficiency virus from  
RT Zaire: nucleotide sequence analysis identifies conserved and variable  
RT domains in the envelope gene";  
RL Gene 52:71-82(1987).  
RN [2]  
RP POST-TRANSLATIONAL MODIFICATIONS, AND FUNCTION.  
RX MEDLINE-88039140; PubMed=3118220;  
RA Guy B., Kleny M.-P., Riviere Y., le Peuch C., Dott K., Girard M.,  
RA Montagnier L., Lecocq J.-P.;  
RT "HIV E/3' orf encodes a phosphorylated GTP-binding protein resembling  
RT an oncogene product";  
RL Nature 330:266-269(1987).  
CC -!- FUNCTION: NEF HAS GTPASE, GTP-BINDING AND AUTOPHOSPHORYLATING  
CC ACTIVITIES, IT SEEMS TO DOWN-REGULATE THE CD4(T4) ANTIGEN.  
CC  
CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
CC the European Bioinformatics Institute. There are no restrictions on its  
CC use by non-profit institutions as long as its content is in no way  
CC modified and this statement is not removed. Usage by and for commercial  
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
CC  
CC EMBL; K03458; AAA45381.1; -  
CC PIR; F26192; QQLJZR.  
CC HSP; P03406; 1EFN.  
CC HIV; K03458; NEFS26.  
CC InterPro; IPR001558; F-protein.  
CC Pfam; PF00469; F-protein; 1.  
CC AIDS; Myristate; GTP-binding.  
CC LIPID 2  
CC MYRISTATE.  
CC SEQUENCE 212 AA; 24415 MW; EC095A1B263047CF CRC64;

Query Match 54.0%; Score 911; DB 1; Length 212;  
Best Local Similarity 77.4%; Pred. No. 1.3e-72;  
Matches 164; Conservative 23; Mismatches 19; Indels 6; Gaps 2;

QY 1 MGGKWSKSSVVGWPTVRRMRRAEPAADGVGAASRDLEKKGAISSNTAATNAC 55  
DB 1 MGGKWSKSSVVGWPAIRIRIRRTNPAADGVGAASRDLEKKGAISSNTRDTNADC 60  
QY 56 AWLEAQEE-EEVGFPPVTPQVPLRPMTYKAAVDLSHFLKKGGLGLIHSORRODILW 114  
DB 61 AWLEAQEESEVGFPPVTPQVPLRPMTYKAAVDLSHFLKKGGLGLIHSORRODILW 120  
QY 115 YHTGQYTPDQNYTPGPGVRYPLTFGWCYKLVPEPKVEANKGENTSLHVPVSLHGM 174  
DB 121 YHTGQYTPDQNYTPGPGIRYPLTFGWCYKLVPEPKVEANKGENTSLHVPVSLHGM 180  
QY 175 DPEREVLEWRDRLAFHVAARELHPEYKNC 206  
DB 181 DTEREVLEWRDRLAFHVAARELHPEYKNC 212

RESULT 16  
ID\_NEF\_HV1ND STANDARD; PRT; 207 AA.  
AC P18801;  
DT 01-NOV-1990 (Rel. 16, Created)  
DT 01-NOV-1990 (Rel. 16, Last sequence update)  
DT 16-OCT-2001 (Rel. 40, Last annotation update)  
DE Negative factor (F-protein) (27 kDa protein) (3'ORF).  
GN NEF.  
OS Human immunodeficiency virus type 1 (NDK isolate) (HIV-1).  
OC Viruses; Retroviral viruses; Retroviridae; Lentivirus.  
OX NCBI\_TaxID=11695;  
RN [1]  
RP SEQUENCE FROM N.A.  
RX MEDLINE-90034200; PubMed=2806917;  
RA Spire B., Sire J., Zachar V., Rey F., Barre-Sinoussi F., Galibert F.,  
RA Hampe A., Chermann J.C.;  
RT "Nucleotide sequence of HIV1-NDK: a highly cytopathic strain of the  
RT human immunodeficiency virus";  
RL Gene 81:275-284(1989).  
CC -!- FUNCTION: NEF HAS GTPASE, GTP-BINDING AND AUTOPHOSPHORYLATING  
CC ACTIVITIES, IT SEEMS TO DOWN-REGULATE THE CD4(T4) ANTIGEN.  
CC -!- MISCELLANEOUS: NDK, ISOLATED FROM A ZAIRIAN PATIENT AFFECTED WITH  
CC AIDS, AND IS A HIGHLY CYTOPATHOGENIC STRAIN.  
CC  
CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
CC the European Bioinformatics Institute. There are no restrictions on its  
CC use by non-profit institutions as long as its content is in no way  
CC modified and this statement is not removed. Usage by and for commercial  
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
CC  
CC EMBL; M27323; AAA44874.1; -  
CC PIR; JQ0068; QQLJND.  
CC HSP; P03406; 1EFN.  
CC HIV; M27323; NEFSNDK.  
CC InterPro; IPR001558; F-protein.  
CC Pfam; PF00469; F-protein; 1.  
CC AIDS; Myristate; GTP-binding.  
CC LIPID 2  
CC MYRISTATE (BY SIMILARITY).  
CC SEQUENCE 207 AA; 23748 MW; 09036C2F81D45D5E CRC64;

Query Match 53.5%; Score 903.5; DB 1; Length 207;  
Best Local Similarity 77.3%; Pred. No. 5.9e-72;  
Matches 160; Conservative 25; Mismatches 21; Indels 1; Gaps 1;

QY 1 MGGKWSKSSVVGWPTVRRMRRAEPAADGVGAASRDLEKKGAISSNTAATNACAWLEA 60  
DB 1 MGGKWSKSSVVGWPAIRIRIRRTNPAADGVGAASRDLEKKGAISSNTASTNADCAWLEA 60  
QY 61 QEE-EEVGFPPVTPQVPLRPMTYKAAVDLSHFLKKGGLGLIHSORRODILWYHTQG 119  
DB 61 QEESEVGFPPVTPQVPLRPMTYKAAVDLSHFLKKGGLGLIHSORRODILWYHTQG 120



```
QY 120 YFPDMQNTPGGVRYPLTFGWCYKLVPEPKVEEANKGENTSLHPVSLHGMDDPERE 179
D 121 IFPDQNTTPGGIRYPLTFGWCYKLVPEPKVEEANKGENTSLHPVSLHGMDDPERE 180
QY 180 VLEWRFDSRLAFHVAHRELHPEYFKN 206
D 181 VLMWRFNSRLAEHKAHRELHPEYFKDC 207

RESULT 17
NEF_HV1U4
ID NEF_HV1U4 STANDARD; PRT: 205 AA.
AC P24741;
DT 01-MAR-1992 (Rel. 21, Created)
DT 01-MAR-1992 (Rel. 21, Last sequence update)
DT 16-OCT-2001 (Rel. 40, Last annotation update)
DE Negative factor (F-protein) (27 kDa protein) (3'ORF).
GN NEF.
OS Human immunodeficiency virus type 1 (strain Ugandan / isolate U455)
OS (HIV-1).
OS Viruses; Retrovirdae; Lentivirus.
OC NCBI_TaxID=11703;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=91090981; PubMed=2265025;
RA Oram J.D., Downing R.G., Roff M., Clegg J.C.S., Serwadda D.,
RA Carswell J.W.;
RT "Nucleotide sequence of a Ugandan HIV-1 provirus reveals genetic
RT diversity from other HIV-1 isolates."
RL AIDS Res. Hum. Retroviruses 6:1073-1078(1990).
CC - FUNCTION: NEF HAS GTPASE, GTP-BINDING AND AUTOPHOSPHORYLATING
CC ACTIVITIES, IT SEEMS TO DOWN-REGULATE THE CD4(T4) ANTIGEN.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL Outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL: M62320; AAA75023.1;
DR HSP: P03406; IEPN.
DR InterPro: IPR001558; F-protein.
DR Pfam: PF00469; F-protein; 1.
KW AIDS; Myristate; GTP-binding.
FT LIPID 2
SQ SEQUENCE 205 AA; 23253 MW; 869AB03E6E7893C4 CRC64;

Query Match 51.7%; Score 872; DB 1; Length 205;
Best Local Similarity 75.1%; Pred. No. 3.3e-69;
Matches 154; Conservative 24; Mismatches 27; Indels 0; Gaps 0;

QY 1 MGKWSKSSVWGVTVRMRRAEPAADGVGAASRDLEKHGATTSNTAATNAACAWLEA 60
D 1 MGKWSKSRVEPEVKRMRTPAARKGVAGVQDLKYGAVTSNTSNTNASCWLEA 60

QY 61 QEEVEYGFVTPQVPLRPMYTKAAVDLSHFLKEKGLGSLHSQRQDILDLYHTQY 120
D 61 QEEGVDGFPVPQVPLRPMYTKAAVDLSHFLKEKGLGSLHSQRQDILDLYHTQY 120

QY 121 FPDQNTTPGGVRYPLTFGWCYKLVPEPKVEEANKGENTSLHPVSLHGMDDPERE 180
D 121 FPDQNTTPGGVRYPLTFGWCYKLVPEPKVEEANKGENTSLHPVSLHGMDDPERE 180

QY 181 LEWRFDSRLAFHVAHRELHPEYFKN 205
D 181 LMKWFDSTLAKHRAVELHPEYFKD 205

RESULT 18
NEF_HV1MN
ID NEF_HV1MN STANDARD; PRT: 182 AA.
AC P05856;
DT 01-NOV-1988 (Rel. 09, Created)
DT 01-NOV-1988 (Rel. 09, Last sequence update)
DT 16-OCT-2001 (Rel. 40, Last annotation update)
DE Negative factor (F-protein) (27 kDa protein) (3'ORF).
GN NEF.
OS Human immunodeficiency virus type 1 (MN isolate) (HIV-1).
OS Viruses; Retrovirdae; Lentivirus.
OC NCBI_TaxID=11696;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=88219542; PubMed=3369091;
RA Gurgo C., Guo H.-G., Franchini G., Aldovini A., Collalti E.,
RA Farrell K., Wong-Staal F., Gallo R.C., Reitz M.S. Jr.;
RT "Envelope sequences of two new United States HIV-1 isolates."
RL Virology 164:531-536(1988).
CC - FUNCTION: NEF HAS GTPASE, GTP-BINDING AND AUTOPHOSPHORYLATING
CC ACTIVITIES, IT SEEMS TO DOWN-REGULATE THE CD4(T4) ANTIGEN.
CC - MISCELLANEOUS: THE MN ISOLATE WAS TAKEN FROM A PEDIATRIC AIDS
CC PATIENT IN 1984.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL Outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL: M17449; AAA44858.1;
DR HSP: P03406; IEPN.
DR HIV: M17449; NEFSMN.
DR InterPro: IPR001558; F-protein.
DR Pfam: PF00469; F-protein; 1.
KW AIDS; Myristate; GTP-binding.
FT LIPID 2
SQ SEQUENCE 182 AA; 20632 MW; 2C5B5A6E3EF13B26 CRC64;

Query Match 49.9%; Score 842; DB 1; Length 182;
Best Local Similarity 87.4%; Pred. No. 1.2e-66;
Matches 153; Conservative 11; Mismatches 7; Indels 4; Gaps 2;

QY 1 MGKWSKSSVWGVTVRMRRAEPAADGVGAASRDLEKHGATTSNTAATNAACAW 57
D 1 MGKWSK-RVTGMPTVRMRRAEPAELAADGVGAASRDLEKHGATTSNTAATNADCAW 59

QY 58 LEAQEEYGFVTPQVPLRPMYTKAAVDLSHFLKEKGLGSLHSQRQDILDLYHT 117
D 60 LEAQEEYGFVTPQVPLRPMYTKAAVDLSHFLKEKGLGSLHSQRQDILDLYHT 119

QY 118 QGYFPDMQNTTPGGVRYPLTFGWCYKLVPEPKVEEANKGENTSLHPVSLHGM 172
D 120 QGYFPDMQNTTPGGVRYPLTFGWCYKLVPEPKVEEANKGENTSLHPVSLHGM 174

RESULT 19
NEF_HV1ZH
ID NEF_HV1ZH STANDARD; PRT: 205 AA.
AC P05859;
DT 01-NOV-1988 (Rel. 09, Created)
DT 01-NOV-1988 (Rel. 09, Last sequence update)
DT 16-OCT-2001 (Rel. 40, Last annotation update)
DE Negative factor (F-protein) (27 kDa protein) (3'ORF).
GN NEF.
OS Human immunodeficiency virus type 1 (Zaire H2321 isolate) (HIV-1).
OS Viruses; Retrovirdae; Lentivirus.
OC NCBI_TaxID=11692;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=89228766; PubMed=2713163;
RA Srinivasan A., York D., Butler D. Jr., Jannoun-Nasr R., Getchell J.,
```

```
RA McCormick J., Ou C.Y., Myers G., Smith T., Chen E.;  
RT "Molecular characterization of HIV-1 isolated from a serum collected  
RT in 1976: nucleotide sequence comparison to recent isolates and  
RT generation of hybrid HIV-1";  
RL AIDS Res. Hum. Retroviruses 5:121-129(1989).  
CC -!- FUNCTION: NEF HAS GTPASE, GTP-BINDING AND AUTOPHOSPHORYLATING  
CC ACTIVITIES, IT SEEMS TO DOWN-REGULATE THE CD4(T4) ANTIGEN.  
CC -----  
CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
CC the European Bioinformatics Institute. There are no restrictions on its  
CC use by non-profit institutions as long as its content is in no way  
CC modified and this statement is not removed. Usage by and for commercial  
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/  
CC or send an email to license@isb-sib.ch).  
CC -----  
DR EMBL; M15896; AAB53951.1; -  
DR PIR; B44963; B44963.  
DR HSP; P03406; IEFN.  
DR HIV; M15896; NEFSZ321.  
DR InterPro: IPR001558; F-protein.  
DR Pfam: PF00469; F-protein; 1.  
KW AIDS; Myristate; GTP-binding.  
FT LIPID 2  
SQ SEQUENCE 205 AA; 23306 MW; EA55B18AF412A9D0 CRC64;  
MYRISTATE (BY SIMILARITY).  
-----  
Query Match 49.6%; Score 837; DB 1; Length 205;  
Best Local Similarity 73.2%; Pred. No. 3.8e-66;  
Matches 153; Conservative 23; Mismatches 25; Indels 8; Gaps 3;  
QY 1 MGGKWSKSVGVPTVRRMR---AEPADGVGAASRDLEKKGALTSSNTAATNAACAW 57  
DB 1 MGNKSK-----GWPAVREIRTPAPPAREGVGAASQDLAKHGLSSNTATNPDCAW 56  
QY 58 LEAQEE-EVGFPPVTPQVPLRPMTYKAAVDLSHFLKEKGLGEGLSHRSORRQDILDWYH 116  
DB 57 LEAQEESEEVGFPPVTPQVPLRPMTYKAAVDLSHFLKEKGLGEGLSHRSORRQDILDWYH 116  
QY 117 TGGYFPDMQNTPGCVRYPLTFGCYKLVPEPKVEEANKGENTSLLLHPSVSLHGMDDP 176  
DB 117 TGGFFPDHNYTPGPGTRPLGFCWCFKLVPEPKVEEANKGENTSLLLHPSVSLHGMDDP 176  
QY 177 EREVLWRFDSRLAFHVAHRELHPEYFKN 205  
DB 177 EREVLWKFDDSLARKHLAREHPEYK 205  
RESULT 20  
NEF_HV1MA  
ID NEF_HV1MA STANDARD; PRT; 209 AA.  
AC P04603.  
DT 13-AUG-1987 (Rel. 05, Created)  
DT 13-AUG-1987 (Rel. 05, Last sequence update)  
DT 16-OCT-2001 (Rel. 40, Last annotation update)  
DE Negative factor (F-protein) (27 kDa protein) (3'ORF).  
GN NEF.  
OS Human immunodeficiency virus type 1 (MAL isolate) (HIV-1).  
OC Viruses; Retroviral viruses; Retroviridae; Lentivirus.  
OX NCBI_TaxID=11697;  
RN [1]  
RP SEQUENCE FROM N.A.  
RX MEDLINE=86245056; PubMed=2424612;  
RA Alison M., Wain-Hobson S., Montagnier L., Sonigo P.;  
RT "Genetic variability of the AIDS virus: nucleotide sequence analysis  
RT of two isolates from African patients.";  
RL Cell 46:63-74(1986).  
RN [2]  
RP POST-TRANSLATIONAL MODIFICATIONS, AND FUNCTION.  
RX MEDLINE=88039140; PubMed=3118220;  
RA Guy B., Kieny M.-P., Riviere Y., le Peuch C., Dott K., Girard M.,  
RA Montagnier L., Lecocq J.-P.;  
RT "Hiv F/3' orf encodes a phosphorylated GTP-binding protein resembling  
an oncogene product.";  
RL Nature 330:266-269(1987).  
CC -!- FUNCTION: NEF HAS GTPASE, GTP-BINDING AND AUTOPHOSPHORYLATING  
CC ACTIVITIES, IT SEEMS TO DOWN-REGULATE THE CD4(T4) ANTIGEN.  
CC -----  
CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
CC the European Bioinformatics Institute. There are no restrictions on its  
CC use by non-profit institutions as long as its content is in no way  
CC modified and this statement is not removed. Usage by and for commercial  
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/  
CC or send an email to license@isb-sib.ch).  
CC -----  
DR EMBL; X04415; CAA28017.1; -  
DR EMBL; A07116; CAA00624.1; -  
DR HSP; P03406; IEFN.  
DR HIV; K03456; NEFSMAL.  
DR InterPro: IPR001558; F-protein.  
DR Pfam: PF00469; F-protein; 1.  
KW AIDS; Myristate; GTP-binding.  
FT LIPID 2  
SQ SEQUENCE 209 AA; 23644 MW; DOB30A2442C8CC44 CRC64;  
MYRISTATE.  
-----  
Query Match 49.4%; Score 834.5; DB 1; Length 209;  
Best Local Similarity 70.1%; Pred. No. 6.4e-66;  
Matches 148; Conservative 26; Mismatches 30; Indels 7; Gaps 2;  
QY 1 MGGKWSKSVGVPTVRRMRRAEPAADGV-----AASRDLEKKGALTSSNTAATNAAC 55  
DB 1 MGGKWSKSVGVPTVRRMRRAEPAADGV-----AASRDLEKKGALTSSNTAATNAAC 55  
QY 56 AWLEAQEESEEVGFPPVTPQVPLRPMTYKAAVDLSHFLKEKGLGEGLSHRSORRQDILDWY 115  
DB 61 E---PPEEEVEVGFPPVTPQVPLRPMTYKAAVDLSHFLKEKGLGEGLSHRSORRQDILDWY 118  
QY 116 HTQGYFPDMQNTPGCVRYPLTFGCYKLVPEPKVEEANKGENTSLLLHPSVSLHGMDD 175  
DB 119 HTQGYFPDMQNTPGCVRYPLTFGCYKLVPEPKVEEANKGENTSLLLHPSVSLHGMDD 175  
QY 176 PEREVLWRFDSRLAFHVAHRELHPEYFKN 206  
DB 179 AEREVLWKFDDSLARKHLAREHPEYKDC 209  
RESULT 21  
NEF_SIVCZ  
ID NEF_SIVCZ STANDARD; PRT; 205 AA.  
AC P17664;  
DT 01-AUG-1990 (Rel. 15, Created)  
DT 01-AUG-1990 (Rel. 15, Last sequence update)  
DT 16-OCT-2001 (Rel. 40, Last annotation update)  
DE Negative factor (F-protein) (27 kDa protein) (3'ORF).  
GN NEF.  
OS Chimpanzee immunodeficiency virus (SIV(cpz)) (CIV).  
OC Viruses; Retroviral viruses; Retroviridae; Lentivirus.  
OX NCBI_TaxID=11723;  
RN [1]  
RP SEQUENCE FROM N.A.  
RX MEDLINE=90259077; PubMed=2188136;  
RA Huet T., Cheynier R., Meyerhans A., Roelants G., Wain-Hobson S.;  
RT "Genetic organization of a chimpanzee lentivirus related to HIV-1.";  
RL Nature 345:356-359(1990).  
CC -!- FUNCTION: NEF HAS GTPASE, GTP-BINDING AND AUTOPHOSPHORYLATING  
CC ACTIVITIES, IT SEEMS TO DOWN-REGULATE THE CD4(T4) ANTIGEN.  
CC -----  
CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
CC the European Bioinformatics Institute. There are no restrictions on its  
CC use by non-profit institutions as long as its content is in no way  
CC modified and this statement is not removed. Usage by and for commercial  
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/  
CC or send an email to license@isb-sib.ch).
```

```
CC -----
DR EMBL; X52154; CAA36408.1;
DR PIR; S09991; ASLJIK.
DR HSP; P03406; IEFN.
DR HIV; X52154; NEFSCP2.
DR InterPro; IPR001558; F-protein.
DR Pfam; PF00469; F-protein; 1.
KW AIDS; Myristate; GTP-binding.
FT LIPID 2
SQ SEQUENCE 205 AA; 23850 MW; 21E0A3EC99F1811F CRC64;

Query Match 49.3%; Score 832; DB 1; Length 205;
Best Local Similarity 70.2%; Pred. No. 1e-65;
Matches 144; Conservative 27; Mismatches 34; Indels 0; Gaps 0;

QY 1 MGCKNSKSSVGVGPTVRMRRAEPAADGVGAASRDLEKHGAITSNTATNAACAWLEA 60
DB 1 MGCKNSKSLVGVGPTVRMRRAEPAADGVGAASRDLEKHGAITSNTATNAACAWLEA 60
QY 61 QEEVEGVFPVTPQVPLRPMTYKAADLSHFLKEKGLGSLHSQRQDILDLYIHTQGY 120
DB 61 MDNEEVGFPVTPQVPLRPMTYKAADLSHFLKEKGLGSLHSQRQDILDLYIHTQGF 120
QY 121 FPDQNYTTPGVRYPPLTFGWCYKLVPEPKVEANKGENTSLHPVSLHGMDDPEREV 180
DB 121 FPDQNYTTPGVRYPPLTFGWCYKLVPEPKVEANKGENTSLHPVSLHGMDDPEREV 180
QY 181 LEWRFDSRLAFHHVARELHPEYKNC 205
DB 181 LVWRFDSRLAFHHVARELHPEYKNC 205

RESULT 22
NEF_HV1SC
ID NEF_HV1SC STANDARD; PRT; 239 AA.
AC P03857;
DT 01-NOV-1988 (Rel. 09, Created)
DT 01-NOV-1988 (Rel. 09, Last sequence update)
DT 16-OCT-2001 (Rel. 40, Last annotation update)
DE Negative factor (F-protein) (27 kDa protein) (3'ORF).
NEF.
OS Human immunodeficiency virus type 1 (SC isolate) (HIV-1).
OC Viruses; Retroviral viruses; Retroviridae; Lentivirus.
OX NCBI_TaxID=11702;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=88219542; PubMed=3369091;
RA Guo C., Guo H.-G., Franchini G., Aldovini A., Collalti E.,
RA Farrell K., Wong-Staal F., Gallo R.C., Reitz M.S. Jr.;
RT "Envelope sequences of two new United States HIV-1 isolates.";
RL Virology 164:531-536(1988).
CC -1- FUNCTION: NEF HAS GTPASE, GTP-BINDING AND AUTOPHOSPHORYLATING
CC ACTIVITIES, IT SEEMS TO DOWN-REGULATE THE CD4(T4) ANTIGEN.
CC -1- MISCELLANEOUS: THE SC ISOLATE WAS TAKEN FROM AN ARC PATIENT IN
CC 1984 IN SOUTHERN CALIFORNIA.
CC
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (see http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC
CC EMBL; M17450; AAA5064.1;
CC HSP; P03406; IEFN.
CC HIV; M17450; NEFSCP.
CC InterPro; IPR001558; F-protein.
CC Pfam; PF00469; F-protein; 1.
KW AIDS; Myristate; GTP-binding.
FT LIPID 2
SQ SEQUENCE 239 AA; 26799 MW; 704A17E54763A99B CRC64;

Query Match 48.5%; Score 819; DB 1; Length 239;
Best Local Similarity 75.5%; Pred. No. 1.7e-64;
Matches 157; Conservative 16; Mismatches 31; Indels 4; Gaps 2;

QY 1 MGCKNSKSSVGVGPTVRMRRAEPAADGVGAASRDLEKHGAITSNTATNAACAWLEA 60
DB 1 MGCKNSKSSVGVGPTVRMRRAEPAADGVGAASRDLEKHGAITSNTATNAACAWLEA 60
QY 61 QEEVEGVFPVTPQVPLRPMTYKAADLSHFLKEKGLGSLHSQRQDILDLYIHTQGY 120
DB 61 QEEVEGVFPVTPQVPLRPMTYKAADLSHFLKEKGLGSLHSQRQDILDLYIHTQGY 118
QY 121 FPDQNYTTPGVRYPPLTFGWCYKLVPEPKVEANKGENTSLHPVSLHGMDDPER 178
DB 119 ATSLIGRTTHQGGSDIPLCFGWCYKLVPEPKVEANKGENTSLHPVSLHGMDDPER 178
QY 179 EVLEWRFDSRLAFHHVARELHPEYKNC 206
DB 179 EVLEWRFDSRLAFHHVARELHPEYKNC 206

RESULT 23
NEF_HV1H2
ID NEF_HV1H2 STANDARD; PRT; 123 AA.
AC P04601; O09780;
DT 13-AUG-1987 (Rel. 05, Created)
DT 15-JUL-1999 (Rel. 38, Last sequence update)
DT 16-OCT-2001 (Rel. 40, Last annotation update)
DE Negative factor (F-protein) (27 kDa protein) (3'ORF).
NEF.
OS Human immunodeficiency virus type 1 (HXB2 isolate) (HIV-1).
OC Viruses; Retroviral viruses; Retroviridae; Lentivirus.
OX NCBI_TaxID=11706;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=87299196; PubMed=3040055;
RA Ratner L., Fisher A., Jagodzinski L.L., Mitsuya H., Liou R.-S.,
RA Gallo R.C., Wong-Staal F.;
RT "Complete nucleotide sequences of functional clones of the AIDS
RT virus.";
RL AIDS Res. Hum. Retroviruses 3:57-69(1987).
RN [2]
RP REVISIONS.
RA Ratner L., Fisher A., Jagodzinski L.L., Mitsuya H., Liou R.-S.,
RA Gallo R.C., Wong-Staal F.;
RL Submitted (APR-1997) to the EMBL/GenBank/DBJ databases.
RN [3]
RP POST-TRANSLATIONAL MODIFICATIONS, AND FUNCTION.
RX MEDLINE=88039140; PubMed=3118220;
RA Guy B., Kiely M.-P., Riviere Y., le Peuch C., Dott K., Girard M.,
RA Montagnier L., Lecocq J.-P.;
RT "HIV F/3' orf encodes a phosphorylated GTP-binding protein resembling
RT an oncogene product.";
RL Nature 330:266-269(1987).
CC -1- FUNCTION: NEF HAS GTPASE, GTP-BINDING AND AUTOPHOSPHORYLATING
CC ACTIVITIES, IT SEEMS TO DOWN-REGULATE THE CD4(T4) ANTIGEN.
CC -1- MISCELLANEOUS: IN THIS ISOLATE A MUTATION IN POSITION 124 ADDS A
CC STOP CODON AND THE PROTEIN IS MUCH SHORTER THAN IN OTHER ISOLATES
CC (210 AA).
CC
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (see http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC
CC EMBL; K03455; AAB50263.1;
CC HSP; AF033819; AAC82597.1;
CC HSP; Q70627; 2NEF.
```

DR HIV: K03455; NEFSHXB2.  
 DR InterPro: IPR001558; F-protein.  
 DR Pfam: PF00469; F-protein; 1.  
 KW AIDS; Myristate; GTP-binding; Phosphorylation.  
 FT LIPID 2 MYRISTATE.  
 FT MOD\_RES 15 15 PHOSPHORYLATION (BY PKC).  
 SQ SEQUENCE 123 AA; 13692 MW; B5007753CCD244CF CRC64;

Query Match 38.2%; Score 645; DB 1; Length 123;  
 Best Local Similarity 98.4%; Pred. No. 1.2e-49;  
 Matches 121; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

QY 1 MGKSKSSVVGWPTVVRMRRAEPAADGVGAASRDLEKKGAISSNTAATNAACAWLEA 60  
 DB 1 MGKSKSSVVGWPTVVRMRRAEPAADGVGAASRDLEKKGAISSNTAATNAACAWLEA 60  
 QY 61 QEEVEGFPVTPQVPLRPMTYKAAVDLSHFLKEKGGLEGLIHSORQDILDWIYHTQGY 120  
 DB 61 QEEVEGFPVTPQVPLRPMTYKAAVDLSHFLKEKGGLEGLIHSORQDILDWIYHTQGY 120  
 QY 121 FPD 123  
 DB 121 FPD 123

RESULT 24  
 ID TAT\_HV1B1 STANDARD; PRT; 123 AA.  
 AC P03404;  
 DT 21-JUL-1986 (Rel. 01, Created)  
 DT 13-AUG-1987 (Rel. 05, Last sequence update)  
 DT 16-OCT-2001 (Rel. 40, Last annotation update)  
 DE Negative factor (F-protein) (27 kDa protein) (3'ORF).  
 GN NEF.  
 OS Human immunodeficiency virus type 1 (BH10 isolate) (HIV-1).  
 OC Viruses; Retroviral viruses; Retroviridae; Lentivirus.  
 OX NCBI\_TaxID=11678;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RX MEDLINE=85111123; PubMed=2578615;  
 RA Ratner L., Haseltine W., Patarca R., Livak K.J., Starcich B.R.,  
 RA Josephs S.F., Doran E.R., Rafalski J.A., Whitehorn E.A.,  
 RA Baumeister K., Ivanoff L., Petteway S.R. Jr., Pearson M.L.,  
 RA Lautenberger J.A., Papas T.S., Grayeb J., Chang N.T., Gallo R.C.,  
 RA Wong-Staal F.;  
 RT "Complete nucleotide sequence of the AIDS virus, HTLV-III.";  
 RL Nature 313:277-284(1985).  
 RN [2]  
 RP POST-TRANSLATIONAL MODIFICATIONS, AND FUNCTION.  
 RX MEDLINE=88039140; PubMed=3118220;  
 RA Guy B., Kieny M.-P., Riviere Y., le Peuch C., Dott K., Girard M.,  
 RA Montagnier L., Lecocq J.-P.;  
 RT "HIV F/3' orf encodes a phosphorylated GTP-binding protein resembling  
 RT an oncogene product.";  
 RL Nature 330:266-269(1987).  
 CC -1- FUNCTION: NEF HAS GTPASE, GTP-BINDING AND AUTOPHOSPHORYLATING  
 CC ACTIVITIES. IT SEEMS TO DOWN-REGULATE THE CD4(74) ANTIGEN.  
 CC -1- MISCELLANEOUS: IN THIS ISOLATE A MUTATION IN POSITION 124 ADDS A  
 CC STOP CODON AND THE PROTEIN IS MUCH SHORTER THAN IN OTHER ISOLATES  
 CC (210 AA).  
 CC -----  
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
 CC the European Bioinformatics Institute. There are no restrictions on its  
 CC use by non-profit institutions as long as its content is in no way  
 CC modified and this statement is not removed. Usage by and for commercial  
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announcement/>  
 CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
 CC -----  
 CC EMBL: M15654; AAA44206.1;  
 DR PIR: A04005; ASLJH3.  
 DR HSP; Q70627; 2NEF.

DR HIV: M15654; NEFSBH102.  
 DR InterPro: IPR001558; F-protein.  
 DR Pfam: PF00469; F-protein; 1.  
 KW AIDS; Myristate; GTP-binding.  
 FT LIPID 2 MYRISTATE.  
 SQ SEQUENCE 123 AA; 13606 MW; 0811735345F0EB8B CRC64;

Query Match 37.8%; Score 638; DB 1; Length 123;  
 Best Local Similarity 97.6%; Pred. No. 5.1e-49;  
 Matches 120; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

QY 1 MGKSKSSVVGWPTVVRMRRAEPAADGVGAASRDLEKKGAISSNTAATNAACAWLEA 60  
 DB 1 MGKSKSSVVGWPTVVRMRRAEPAADGVGAASRDLEKKGAISSNTAATNAACAWLEA 60  
 QY 61 QEEVEGFPVTPQVPLRPMTYKAAVDLSHFLKEKGGLEGLIHSORQDILDWIYHTQGY 120  
 DB 61 QEEVEGFPVTPQVPLRPMTYKAAVDLSHFLKEKGGLEGLIHSORQDILDWIYHTQGY 120  
 QY 121 FPD 123  
 DB 121 FPD 123

RESULT 25  
 ID TAT\_HV1B1 STANDARD; PRT; 86 AA.  
 AC P04506;  
 DT 13-AUG-1987 (Rel. 05, Created)  
 DT 13-AUG-1987 (Rel. 05, Last sequence update)  
 DT 16-OCT-2001 (Rel. 40, Last annotation update)  
 DE TAT protein (Transactivating regulatory protein).  
 GN TAT.  
 OS Human immunodeficiency virus type 1 (BH10 isolate) (HIV-1), and  
 OS Human immunodeficiency virus type 1 (HXB3 isolate) (HIV-1).  
 OC Viruses; Retroviral viruses; Retroviridae; Lentivirus.  
 OX NCBI\_TaxID=11678, 11707;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RX STRAIN=ISOLATE BH10;  
 RX MEDLINE=85111123; PubMed=2578615;  
 RA Ratner L., Haseltine W., Patarca R., Livak K.J., Starcich B.R.,  
 RA Josephs S.F., Doran E.R., Rafalski J.A., Whitehorn E.A.,  
 RA Baumeister K., Ivanoff L., Petteway S.R. Jr., Pearson M.L.,  
 RA Lautenberger J.A., Papas T.S., Grayeb J., Chang N.T., Gallo R.C.,  
 RA Wong-Staal F.;  
 RT "Complete nucleotide sequence of the AIDS virus, HTLV-III.";  
 RL Nature 313:277-284(1985).  
 RN [2]  
 RP SEQUENCE FROM N.A.  
 RX STRAIN=ISOLATE HXB3;  
 RX MEDLINE=85228248; PubMed=2988795;  
 RA Crowl R., Ganguly K., Gordon M., Conroy R., Schaber M., Kramer R.,  
 RA Shaw G.M., Wong-Staal F., Reddy E.P.;  
 RT "HTLV-III env gene products synthesized in E. coli are recognized by  
 RT antibodies present in the sera of AIDS patients.";  
 RL Cell 41:979-986(1985).  
 CC -1- FUNCTION: TRANSCRIPTIONAL REGULATOR THAT ACTS BY BINDING TO THE  
 CC TRANS-ACTIVATING RESPONSIVE SEQUENCE (TAR) RNA ELEMENT AND  
 CC ACTIVATES TRANSCRIPTION INITIATION AND/OR ELONGATION FROM THE LTR  
 CC PROMOTER.  
 CC -1- SUBUNIT: BINDS CYCLIN T1 (BY SIMILARITY).  
 CC -1- SUBCELLULAR LOCATION: NUCLEAR; NUCLEOLAR.  
 CC -1- MISCELLANEOUS: THE TWO ISOLATES TAT SEQUENCE ARE IDENTICAL.  
 CC -----  
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
 CC the European Bioinformatics Institute. There are no restrictions on its  
 CC use by non-profit institutions as long as its content is in no way  
 CC modified and this statement is not removed. Usage by and for commercial  
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announcement/>  
 CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
 CC -----

```
CC EMBL; M14100; AAA44676.1; -
DR EMBL; M15654; AAA44199.1; -
DR HIV; M15654; TAT$BH102.
DR HIV; M14100; TAT$HXB3.
DR InterPro; IPR001831; HIV_Tat.
DR Pfam; PF00539; Tat; 1.
DR PRINTS; PR00055; HIVTATDOMAIN.
KW Transcription regulation; Activator; RNA-binding; Nuclear protein;
KW AIDS.
SQ SEQUENCE 86 AA; 9784 MW; 4DD609415FAF9015 CRC64;

Query Match 28.8%; Score 486; DB 1; Length 86;
Best Local Similarity 100.0%; Pred. No. 6.4e-36;
Matches 85; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 209 EPVDPRLPEPKHGPSQPKTACTNCYCKKCCFHCQVCFITKALGISYGRKKRRRRPPQG 268
DB 2 EPVDPRLPEPKHGPSQPKTACTNCYCKKCCFHCQVCFITKALGISYGRKKRRRRPPQG 61
QY 269 SOTHQVSLSKOPTSQSRGDPGPK 293
DB 62 SOTHQVSLSKOPTSQSRGDPGPK 86

RESULT 26
TAT_HV1P
ID TAT_HV1P STANDARD; PRT; 86 AA.
AC 04607;
DT 13-AUG-1987 (Rel. 05, Created)
DT 13-AUG-1987 (Rel. 05, Last sequence update)
DT 16-OCT-2001 (Rel. 40, Last annotation update)
DE TAT protein (Transactivating regulatory protein).
GN TAT.
OS Human immunodeficiency virus type 1 (PV22 isolate) (HIV-1).
OC Viruses; Retroviral viruses; Retroviridae; Lentivirus.
OX NCBI_TaxID=11700;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=8511157; PubMed=2982104;
RA Mesing M.A., Smith D.H., Cabradilla C.D., Benton C.V., Lasky L.A.,
RA Capon D.J.;
RT "Nucleic acid structure and expression of the human
RT AIDS/lymphadenopathy retrovirus.";
RL Nature 313:450-458(1985).
CC -!- FUNCTION: TRANSCRIPTIONAL REGULATOR THAT ACTS BY BINDING TO THE
CC TRANS-ACTIVATING RESPONSIVE SEQUENCE (TAR) RNA ELEMENT AND
CC ACTIVATES TRANSCRIPTION INITIATION AND/OR ELONGATION FROM THE LTR
CC PROMOTER.
CC -!- SUBUNIT: BINDS CYCLIN T1 (BY SIMILARITY).
CC -!- SUBCELLULAR LOCATION: NUCLEAR; NUCLEOLAR.
CC [1]
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
CC EMBL; M14100; AAA44999.1; -
CC PIR; A04017; TNLJ12.
CC HIV; M1840; TAT$PCV12.
CC InterPro; IPR001831; HIV_Tat.
CC Pfam; PF00539; Tat; 1.
CC PRINTS; PR00055; HIVTATDOMAIN.
KW Transcription regulation; Activator; RNA-binding; Nuclear protein;
KW AIDS.
SQ SEQUENCE 86 AA; 9794 MW; 4DD5C6415FAF9015 CRC64;
```

Query Match 28.5%; Score 481; DB 1; Length 86;

```
Best Local Similarity 98.8%; Pred. No. 1.7e-35;
Matches 84; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 209 EPVDPRLPEPKHGPSQPKTACTNCYCKKCCFHCQVCFITKALGISYGRKKRRRRPPQG 268
DB 2 EPVDPRLPEPKHGPSQPKTACTNCYCKKCCFHCQVCFITKALGISYGRKKRRRRPPQG 61
QY 269 SOTHQVSLSKOPTSQSRGDPGPK 293
DB 62 SOTHQVSLSKOPTSQSRGDPGPK 86

RESULT 27
TAT_HV12
ID TAT_HV12 STANDARD; PRT; 86 AA.
AC 04326;
DT 20-MAR-1987 (Rel. 04, Created)
DT 13-AUG-1987 (Rel. 05, Last sequence update)
DT 16-OCT-2001 (Rel. 40, Last annotation update)
DE TAT protein (Transactivating regulatory protein).
GN TAT.
OS Human immunodeficiency virus type 1 (clone 12) (HIV-1).
OC Viruses; Retroviral viruses; Retroviridae; Lentivirus.
OX NCBI_TaxID=11679;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=86177573; PubMed=3008154;
RA Arya S.K., Gallo R.C.;
RT "Three novel genes of human T-lymphotropic virus type III: immune
RT reactivity of their products with sera from acquired immune
RT deficiency syndrome patients.";
RL Proc. Natl. Acad. Sci. U.S.A. 83:2209-2213(1986).
CC -!- FUNCTION: TRANSCRIPTIONAL REGULATOR THAT ACTS BY BINDING TO THE
CC TRANS-ACTIVATING RESPONSIVE SEQUENCE (TAR) RNA ELEMENT AND
CC ACTIVATES TRANSCRIPTION INITIATION AND/OR ELONGATION FROM THE LTR
CC PROMOTER.
CC -!- SUBUNIT: BINDS CYCLIN T1 (BY SIMILARITY).
CC -!- SUBCELLULAR LOCATION: NUCLEAR; NUCLEOLAR.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
CC EMBL; M1840; AAA44999.1; -
CC PIR; A04017; TNLJ12.
CC HIV; M1840; TAT$PCV12.
CC InterPro; IPR001831; HIV_Tat.
CC Pfam; PF00539; Tat; 1.
CC PRINTS; PR00055; HIVTATDOMAIN.
KW Transcription regulation; Activator; RNA-binding; Nuclear protein;
KW AIDS.
SQ SEQUENCE 86 AA; 9758 MW; 4DD609414FBE9115 CRC64;

Query Match 28.3%; Score 478; DB 1; Length 86;
Best Local Similarity 98.8%; Pred. No. 3.2e-35;
Matches 84; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 209 EPVDPRLPEPKHGPSQPKTACTNCYCKKCCFHCQVCFITKALGISYGRKKRRRRPPQG 268
DB 2 EPVDPRLPEPKHGPSQPKTACTNCYCKKCCFHCQVCFITKALGISYGRKKRRRRPPQG 61
QY 269 SOTHQVSLSKOPTSQSRGDPGPK 293
DB 62 SOTHQVSLSKOPTSQSRGDPGPK 86

RESULT 28
TAT_HV1B
```

```

ID AC P04610; STANDARD; PRT; 86 AA.
DT 13-AUG-1987 (Rel. 05, Created)
DT 13-AUG-1987 (Rel. 05, Last sequence update)
DT 16-OCT-2001 (Rel. 40, Last annotation update)
DE TAT protein (Transactivating regulatory protein).
GN TAT.
OS Human immunodeficiency virus type 1 (BRU isolate) (HIV-1).
OC Viruses; Retrovirdae; Retroviridae; Lentivirus.
OX NCBI_TaxID=11686;
RN [1]
RP SEQUENCE FROM N.A. (CLONE PNL4-3).
RX MEDLINE=8509333; PubMed=2981635;
RA Wain-Hobson S., Sonigo P., Danos O., Cole S., Alizon M.;
RT "Nucleotide sequence of the AIDS virus, LAV.";
RL Cell 40:9-17(1985).
RN [2]
RP SEQUENCE FROM N.A. (CLONE PNL4-3).
RA Buckler C.E., Buckler-White A.J., Willey R.L., McCoy J.;
RL Submitted (JUN-1988) to the EMBL/GenBank/DBJ databases
CC -!- FUNCTION: TRANSCRIPTIONAL REGULATOR THAT ACTS BY BINDING TO THE
CC TRANS-ACTIVATING RESPONSE SEQUENCE (TAR) RNA ELEMENT AND
CC ACTIVATES TRANSCRIPTION INITIATION AND/OR ELONGATION FROM THE LTR
CC PROMOTER.
CC -!- SUBUNIT: BINDS CYCLIN T1 (BY SIMILARITY).
CC -!- SUBCELLULAR LOCATION: NUCLEAR; NUCLEOLAR.
CC
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC
CC EMBL; K02013; AA059745.1; -.
CC EMBL; M19921; AAA44985.1; -.
CC HIV; K02013; TATSBRU.
CC HIV; M19921; TATSBL43.
CC InterPro: IPR001831; HIV_Tat.
CC Pfam: PF00539; Tat; 1.
CC PRINTS: PR00055; HIVTATDOMAIN.
CC Transcription regulation; Activator; RNA-binding; Nuclear protein;
CC AIDS.
CC
CC VARIANT 24 24 T -> N (IN CLONE PNL4-3).
CC VARIANT 29 39 T -> M (IN CLONE PNL4-3).
CC VARIANT 58 61 PPQG -> AHON (IN CLONE PNL4-3).
CC VARIANT 67 67 V -> A (IN CLONE PNL4-3).
CC VARIANT 77 77 P -> S (IN CLONE PNL4-3).
CC SEQUENCE 86 AA; 9769 MW; 9B1BA4915FAF8A14 CRC64;

Query Match 27.8%; Score 470; DB 1; Length 86;
Best Local Similarity 96.5%; Pred. No. 1.6e-34;
Matches 82; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

QY 209 EPVDRLEPKHPGSPKPTACTNCYCKCCFHCQVCFTKALGISYGRKKRRRPPG 268
DB 2 EPVDRLEPKHPGSPKPTACTTCYCKCCFHCQVCFTKALGISYGRKKRRRPPG 61
QY 269 SOTHQVSLSKQPTSSRSGDTPGPK 293
DB 62 SOTHQVSLSKQPTSSRSGDTPGPK 86

RESULT 29
TAT_HVIRH
ID TAT_HVIRH STANDARD; PRT; 102 AA.
AC P05908;
DT 01-NOV-1988 (Rel. 09, Created)
DT 01-NOV-1988 (Rel. 09, Last sequence update)
DT 16-OCT-2001 (Rel. 40, Last annotation update)
DE TAT protein (Transactivating regulatory protein).
```

```

GN OS
OS Human immunodeficiency virus type 1 (RF/HAT isolate) (HIV-1).
OC Viruses; Retrovirdae; Retroviridae; Lentivirus.
OX NCBI_TaxID=11701;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=86218077; PubMed=2423250;
RA Starcich B.R., Hahn B.H., Shaw G.M., McNeely P.D., Modrow S.,
RA Wolf H., Parks E.S., Parks W.P., Josephs S.F., Gallo R.C.,
RA Wong-Staal F.;
RT "Identification and characterization of conserved and variable
RT regions in the envelope gene of HTLV-III/LAV, the retrovirus of
RT AIDS.";
RL Cell 45:637-648(1986).
CC -!- FUNCTION: TRANSCRIPTIONAL REGULATOR THAT ACTS BY BINDING TO THE
CC TRANS-ACTIVATING RESPONSE SEQUENCE (TAR) RNA ELEMENT AND
CC ACTIVATES TRANSCRIPTION INITIATION AND/OR ELONGATION FROM THE LTR
CC PROMOTER.
CC -!- SUBUNIT: BINDS CYCLIN T1 (BY SIMILARITY).
CC -!- SUBCELLULAR LOCATION: NUCLEAR; NUCLEOLAR.
CC
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC
CC EMBL; M17451; AAA45050.1; -.
CC HIV; M17451; TATSBRF.
CC InterPro: IPR001831; HIV_Tat.
CC Pfam: PF00539; Tat; 1.
CC PRINTS: PR00055; HIVTATDOMAIN.
CC Transcription regulation; Activator; RNA-binding; Nuclear protein;
CC AIDS.
CC SEQUENCE 102 AA; 11538 MW; 3EACFB843C5195BD CRC64;

Query Match 27.3%; Score 461; DB 1; Length 102;
Best Local Similarity 91.9%; Pred. No. 1.2e-33;
Matches 79; Conservative 3; Mismatches 4; Indels 0; Gaps 0;

QY 209 EPVDRLEPKHPGSPKPTACTNCYCKCCFHCQVCFTKALGISYGRKKRRRPPG 268
DB 2 EPVDRLEPKHPGSPKPTACTNCYCKCCFHCQVCFTKALGISYGRKKRRRPPG 61
QY 269 SOTHQVSLSKQPTSSRSGDTPGPK 294
DB 62 SOTHQVSLSKQPTSSRSGDTPGPK 87

RESULT 30
TAT_HVIRH
ID TAT_HVIRH STANDARD; PRT; 86 AA.
AC P04608; O09778;
DT 13-AUG-1987 (Rel. 05, Created)
DT 15-JUL-1999 (Rel. 38, Last sequence update)
DT 16-OCT-2001 (Rel. 40, Last annotation update)
DE TAT protein (Transactivating regulatory protein).
GN TAT.
OS Human immunodeficiency virus type 1 (HXB2 isolate) (HIV-1).
OC Viruses; Retrovirdae; Retroviridae; Lentivirus.
OX NCBI_TaxID=11706;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=87299196; PubMed=3040055;
RA Ratner L., Fisher A., Jagodzinski L.L., Mitsuya H., Liou R.-S.,
RA Gallo R.C., Wong-Staal F.;
RT "Complete nucleotide sequences of functional clones of the AIDS
RT virus.";
RL AIDS Res. Hum. Retroviruses 3:57-69(1987).
RN [2]
```

```

RP REVISIONS.
RA Ratner L., Fisher A., Jagodzinski L.L., Mitsuya H., Liou R.-S.,
RA Gallo R.C., Wong-Staal F.;
RL Submitted (APR-1997) to the EMBL/GenBank/DBJ databases.
CC -!- FUNCTION: TRANSCRIPTIONAL REGULATOR THAT ACTS BY BINDING TO THE
CC TRANS-ACTIVATING RESPONSIVE SEQUENCE (TAR) RNA ELEMENT AND
CC ACTIVATES TRANSCRIPTION INITIATION AND/OR ELONGATION FROM THE LTR
CC PROMOTER.
CC -!- SUBUNIT: BINDS CYCLIN T1 (BY SIMILARITY).
CC -!- SUBCELLULAR LOCATION: NUCLEAR; NUCLEOLAR.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL; K03455; AAB50256.1; -.
DR EMBL; AF033819; AAC82591.1; -.
DR HIV; K03455; TAT$HX82.
DR InterPro; IPR001831; HIV_Tat.
DR Pfam; PF00539; Tat; 1.
DR PRINTS; PR00055; HIVTATDOMAIN.
KW Transcription regulation; Activator; RNA-binding; Nuclear protein;
KW AIDS.
SQ SEQUENCE 86 AA; 9837 MW; 4DDC56D979769115 CRC64;

```

Query Match 26.9%; Score 454; DB 1; Length 86;  
 Best Local Similarity 94.1%; Pred. No. 4e-33;  
 Matches 80; Conservative 0; Mismatches 5; Indels 0; Gaps 0;

```

QY 209 EPVDPRLPEPKHGPSQPKTACTNCTCYKCCFHCQVCFTKALGISYGRKKRRRPPQG 268
   |||||
DB 2 EPVDPRLPEPKHGPSQPKTACTNCTCYKCCFHCQVCFTKALGISYGRKKRRRRAHQN 61
   |||||
QY 269 SOTHQVSLSKQPTQSQRGDTGPKE 293
   |||||
DB 62 SOTHQVSLSKQPTQSQRGDTGPKE 86

```

Search completed: August 26, 2002, 08:15:41  
 Job time: 345 sec

GenCore version 4.5  
Copyright (c) 1993 - 2000 Compugen Ltd.

# OM protein - protein search, using sw model

Run on: August 26, 2002, 08:08:32 ; Search time 86.53 Seconds  
(without alignments)  
603.773 Million cell updates/sec

Title: US-09-509-239-13  
Perfect score: 1688  
Sequence: 1 MGCKWSSVGVGPTVRER.....QSRGDTGPKETSGHHHHH 302

Scoring table:  
BLOSUM62  
Gapop 10.0 , Gapext 0.5

Searched: 562222 seqs, 172994929 residues

Total number of hits satisfying chosen parameters: 562222

Minimum DB seq length: 0  
Maximum DB seq length: 2000000000  
Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

Database : SPTREMBL\_19.\*  
1: sp\_archaea.\*  
2: sp\_bacteria.\*  
3: sp\_fungi.\*  
4: sp\_human.\*  
5: sp\_invertebrate.\*  
6: sp\_mammal.\*  
7: sp\_mhc.\*  
8: sp\_organelle.\*  
9: sp\_phage.\*  
10: sp\_plant.\*  
11: sp\_rodent.\*  
12: sp\_virus.\*  
13: sp\_vertebrate.\*  
14: sp\_unclassified.\*  
15: sp\_rvirus.\*  
16: sp\_bacteriap.\*  
17: sp\_archaeap.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

## SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	1129	66.9	206	15 Q9WM24	Q9wm24 human immun
2	1124	66.6	206	15 Q9WLM4	Q9wlm4 human immun
3	1119	66.3	206	15 Q40177	Q40177 human immun
4	1115	66.1	206	15 Q90VU7	Q90vu7 human immun
5	1108	65.6	206	15 Q9PXM9	Q9pxw9 human immun
6	1105	65.5	206	15 Q9WML6	Q9wml6 human immun
7	1104	65.4	206	15 Q9WM30	Q9wm30 human immun
8	1101	65.2	206	15 Q9PXM8	Q9pxw8 human immun
9	1096	64.9	206	15 Q90179	Q90179 human immun
10	1096	64.9	206	15 Q95588	Q95588 aids-associ
11	1096	64.9	206	15 Q78244	Q78244 human immun
12	1092	64.7	206	15 Q9WLM7	Q9wlm7 human immun
13	1090	64.6	206	15 Q74905	Q74905 human immun
14	1087	64.4	206	15 Q89561	Q89561 human immun
15	1085	64.3	206	15 Q74913	Q74913 human immun
16	1080	64.0	206	15 Q74917	Q74917 human immun

17	1079	63.9	206	15 Q74914	Q74914 human immun
18	1079	63.9	206	15 Q74915	Q74915 human immun
19	1077	63.8	206	15 Q74909	Q74909 human immun
20	1076	63.7	202	15 Q9QPN3	Q9qpn3 human immun
21	1072	63.5	206	15 Q9Q596	Q9q596 human immun
22	1072	63.5	206	15 Q9Q595	Q9q595 human immun
23	1071	63.4	208	15 Q9YU09	Q9yyu9 human immun
24	1068	63.3	206	15 Q9W7X3	Q9w7x3 human immun
25	1065	63.1	206	15 Q9DQ01	Q9dqu1 human immun
26	1064	63.0	206	15 Q9W7U0	Q9w7u0 human immun
27	1062	62.9	206	15 Q93010	Q93010 human immun
28	1061	62.9	206	15 Q93012	Q93012 human immun
29	1060	62.8	206	15 Q9Q597	Q9q597 human immun
30	1057	62.6	206	15 Q9W7X2	Q9w7x2 human immun
31	1054	62.4	206	15 Q9Q594	Q9q594 human immun
32	1053	62.4	206	15 Q9Q5A3	Q9q5a3 human immun
33	1053	62.4	206	15 Q9DQF9	Q9dqt9 human immun
34	1052	62.3	205	15 Q74823	Q74823 human immun
35	1051	62.3	206	15 Q74824	Q74824 human immun
36	1050	62.2	206	15 Q9QR05	Q9qrd5 human immun
37	1050	62.2	206	15 Q9Q599	Q9q599 human immun
38	1049	62.1	206	15 Q89630	Q89630 human immun
39	1049	62.1	206	15 Q71956	Q71956 human immun
40	1048	62.1	206	15 Q9Q5A8	Q9q5a8 human immun
41	1048	62.1	206	15 Q72421	Q72421 human immun
42	1047	62.0	206	15 Q89537	Q89537 human immun
43	1047	62.0	206	15 Q9Q5D3	Q9q5d3 human immun
44	1047	62.0	206	15 Q72985	Q72985 human immun
45	1047	62.0	210	15 P90275	P90275 human immun

## ALIGNMENTS

RESULT 1  
Q9WM24 PRELIMINARY; PRT: 206 AA.  
ID Q9WM24;  
AC Q9WM24;  
DT 01-NOV-1999 (TREMBLrel. 12, Created)  
DT 01-NOV-1999 (TREMBLrel. 12, Last sequence update)  
DT 01-DEC-2001 (TREMBLrel. 19, Last annotation update)  
DE NEGATIVE FACTOR (F-PROTEIN) (27 KDA PROTEIN).  
GN NEF.  
OS Human immunodeficiency virus type 1.  
OC Viruses; Retroid viruses; Retroviridae; Lentivirus.  
OC NCBI\_TaxID=11676;  
RN [1]  
RP SEQUENCE FROM N.A.  
RC STRAIN=12-IT;  
RA Visco Comandini U., Yun Z., Paganelli R., Orlandi P., Salotti A.,  
RA Johansson B., Vahine A., Sonnerborg A.;  
RT "HIV-1 nef mutations and clinical long-term non progression: a  
RT molecular epidemiology study."  
RL Submitted (JUN-1997) to the EMBL/GenBank/DBJ databases.  
CC -I- SIMILARITY: NEF HAS GTPASE, GTP-BINDING AND AUTOPHOSPHORYLATING  
CC ACTIVITIES; IT SEEMS TO DOWN-REGULATE THE CD4(T4) ANTIGEN.  
DR EMBL; AF011480; RAD01458.1;  
DR HSSP; P03406; 1EFN.  
DR InterPro: IPR001558; F-protein.  
DR Pfam: PF00469; F-protein; 1.  
KW Aids; GTP-binding; Myristate.  
SQ SEQUENCE 206 AA; 23328 MW; FCC69458158FIA03 CRC64;

Query Match 66.9%; Score 1129; DB 15; Length 206;  
Best Local Similarity 99.5%; Pred. No. 8.7e-95;  
Matches 205; Conservative 1; Mismatches 0; Indels 0; Gaps 0;  
QY 1 MGCKWSSVGVGPTVRERMRRAEPAADGVGAASRDLEKHGAISSNTAATNAACAWLEA 60  
|||||  
Db 1 MGCKWSSVGVGPTVRERMRRAEPAADGVGAASRDLEKHGAISSNTAATNAACAWLEA 60  
|||||  
QY 61 QEEEEVGFPVTPQVPLRPMTYKAAVDLSHFLKEKGGLEGHSQRQDILDLYHTQGY 120



```

Db 61 QEEVEGFPVTPQVPLRPMTYKAADVLSHFLKEKGLIHSQRQDILDLWYHTQY 120
QY 121 FPDWQNTPGVRYPLTFCWCYKLVPEPDKVEEANKGENTSLHPVSLHGMDPDERV 180
Db 121 FPDWQNTPGVRYPLTFCWCYKLVPEPDKVEEANKGENTSLHPVSLHGMDPDERV 180
QY 181 LEWRDLSRLAFHFAHVAHELHPEYFKNC 206
Db 181 LEWRDLSRLAFHFAHVAHELHPEYFKNC 206

RESULT 2
Q9WLM4
ID Q9WLM4 PRELIMINARY; PRT; 206 AA.
AC Q9WLM4;
DT 01-NOV-1999 (TrEMBLrel. 12, Created)
DT 01-NOV-1999 (TrEMBLrel. 12, Last sequence update)
DT 01-DEC-2001 (TrEMBLrel. 19, Last annotation update)
DE NEGATIVE FACTOR (F-PROTEIN) (27 KDA PROTEIN).
GN NEF.
OS Human immunodeficiency virus type 1.
OC Viruses; Retrovirdae; Retroviridae; Lentivirus.
OX NCBI_TaxID=11676;
RN [1]
RC STRAIN=21-SW;
RA Visco Comandini U., Yun Z., Paganelli R., Orlandi P., Salotti A.,
RA Johansson B., Vahine A., Sonnerborg A.;
RT "Hiv-1 nef mutations and clinical long-term non progression: a
RT molecular epidemiology study.";
RL Submitted (FEB-1998) to the EMBL/GenBank/DBJ databases.
CC -!- SIMILARITY: NEF HAS GTPASE, GTP-BINDING AND AUTOPHOSPHORYLATING
CC ACTIVITIES, IT SEEMS TO DOWN-REGULATE THE CD4(T4) ANTIGEN.
DR EMBL; AF047087; AD02461.1;
DR HSP; P03406; IEFN.
DR InterPro; IPR001558; F-protein.
DR Pfam; PF00469; F-protein; 1.
KW Aids; Gtp-binding; Myristate.
SQ SEQUENCE 206 AA; 23341 MW; FCD22BICEB55BB9 CRC64;

Query Match 66.6%; Score 1124; DB 15; Length 206;
Best Local Similarity 99.0%; Pred. No. 2.5e-94;
Matches 204; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

QY 1 MGGKWSKSSVGVWPTVRRMRRAEPAADGVGAASRDLEKHGATSSNTAATNAACAWLEA 60
Db 1 MGGKWSKSSVGVWPTVRRMRRAEPAADGVGAASRDLEKHGATSSNTAATNAACAWLEA 60
QY 61 QEEVEGFPVTPQVPLRPMTYKAADVLSHFLKEKGLIHSQRQDILDLWYHTQY 120
Db 61 QEEVEGFPVTPQVPLRPMTYKAADVLSHFLKEKGLIHSQRQDILDLWYHTQY 120
QY 121 FPDWQNTPGVRYPLTFCWCYKLVPEPDKVEEANKGENTSLHPVSLHGMDPDERV 180
Db 121 FPDWQNTPGVRYPLTFCWCYKLVPEPDKVEEANKGENTSLHPVSLHGMDPDERV 180
QY 181 LEWRDLSRLAFHFAHVAHELHPEYFKNC 206
Db 181 LEWRDLSRLAFHFAHVAHELHPEYFKNC 206

RESULT 3
Q40177
ID Q40177 PRELIMINARY; PRT; 206 AA.
AC Q40177;
DT 01-JAN-1998 (TrEMBLrel. 05, Created)
DT 01-JAN-1998 (TrEMBLrel. 05, Last sequence update)
DT 01-DEC-2001 (TrEMBLrel. 19, Last annotation update)
DE NEGATIVE FACTOR (F-PROTEIN) (27 KDA PROTEIN).
GN NEF.
OS Human immunodeficiency virus type 1.

```

```

OC Viruses; Retrovirdae; Retroviridae; Lentivirus.
OX NCBI_TaxID=11676;
RN [1]
RC STRAIN=PNL4-3;
RA Fang G., Weiser B., Visosky A., Burger H.;
RT "Constructing full-length chimeric HIV-1 molecular clones by PCR-
RT mediated recombination.";
RL Submitted (MAY-1997) to the EMBL/GenBank/DBJ databases.
CC -!- SIMILARITY: NEF HAS GTPASE, GTP-BINDING AND AUTOPHOSPHORYLATING
CC ACTIVITIES, IT SEEMS TO DOWN-REGULATE THE CD4(T4) ANTIGEN.
DR EMBL; AF003887; AAB64279.1;
DR HSP; P03406; IEFN.
DR InterPro; IPR001558; F-protein.
DR Pfam; PF00469; F-protein; 1.
KW Aids; Gtp-binding; Myristate.
SQ SEQUENCE 206 AA; 23428 MW; 9F94AEB9CAFD6CF6 CRC64;

Query Match 66.3%; Score 1119; DB 15; Length 206;
Best Local Similarity 98.5%; Pred. No. 7e-94;
Matches 203; Conservative 1; Mismatches 2; Indels 0; Gaps 0;

QY 1 MGGKWSKSSVGVWPTVRRMRRAEPAADGVGAASRDLEKHGATSSNTAATNAACAWLEA 60
Db 1 MGGKWSKSSVGVWPTVRRMRRAEPAADGVGAASRDLEKHGATSSNTAATNAACAWLEA 60
QY 61 QEEVEGFPVTPQVPLRPMTYKAADVLSHFLKEKGLIHSQRQDILDLWYHTQY 120
Db 61 QEEVEGFPVTPQVPLRPMTYKAADVLSHFLKEKGLIHSQRQDILDLWYHTQY 120
QY 121 FPDWQNTPGVRYPLTFCWCYKLVPEPDKVEEANKGENTSLHPVSLHGMDPDERV 180
Db 121 FPDWQNTPGVRYPLTFCWCYKLVPEPDKVEEANKGENTSLHPVSLHGMDPDERV 180
QY 181 LEWRDLSRLAFHFAHVAHELHPEYFKNC 206
Db 181 LEWRDLSRLAFHFAHVAHELHPEYFKNC 206

RESULT 4
Q90VU7
ID Q90VU7 PRELIMINARY; PRT; 206 AA.
AC Q90VU7;
DT 01-DEC-2001 (TrEMBLrel. 19, Created)
DT 01-DEC-2001 (TrEMBLrel. 19, Last sequence update)
DT 01-DEC-2001 (TrEMBLrel. 19, Last annotation update)
DE NEF PROTEIN.
GN NEF.
OS Human immunodeficiency virus type 1.
OC Viruses; Retrovirdae; Retroviridae; Lentivirus.
OX NCBI_TaxID=11676;
RN [1]
RC STRAIN=NL4-3;
RA MEDLINE=96036482; PubMed=7483282;
RA Salminen M.O., Koch C., Sanders-Buell E., Ehrenberg P.K.,
RA Michael N.L., Carr J.K., Burke D.S., McCutchan F.E.;
RT "Recovery of virtually full-length HIV-1 provirus of diverse subtypes
RT from primary virus cultures using the polymerase chain reaction.";
RL Virology 213:80-86(1995).
RN [2]
RC SEQUENCE FROM N.A.
RC STRAIN=NL4-3;
RX MEDLINE=86281827; PubMed=3016298;
RA Adachi A., Gendelman H.E., Koenig S., Folks T., Willey R., Rabson A.,
RA Martin M.A.;
RT "Production of acquired immunodeficiency syndrome-associated
RT retrovirus in human and nonhuman cells transfected with an infectious
RT molecular clone.";
RL J. Virol. 59:284-291(1985).
DR EMBL; U26942; AAB60579.1;
SQ SEQUENCE 206 AA; 23367 MW; 65AF3B6184DC2FE7 CRC64;

```

Query Match 66.1%; Score 1115; DB 15; Length 206;  
Best Local Similarity 98.1%; Pred. No. 1.6e-93;  
Matches 202; Conservative 1; Mismatches 3; Indels 0; Gaps 0;

QY 1 MGKWSKSSVVGWPTVRRMRRAEPAADGVGAASRDLEKKGGLHSGRRQDILDLWIYHTGY 60  
DB 1 MGKWSKSSVIGWPAVRRMRRAEPAADGVGAASRDLEKKGGLHSGRRQDILDLWIYHTGY 60

QY 61 QEEEEVGFPVTPQVPLRPMTYKAADVLSHFLKEKGLHSGRRQDILDLWIYHTGY 120  
DB 61 QEEEEVGFPVTPQVPLRPMTYKAADVLSHFLKEKGLHSGRRQDILDLWIYHTGY 120

QY 121 FPDQNTYTPGPGVRYPLTFGWCYKLVPEPDKVEEANKGENTSLHHPVSLHGMDDPEREV 180  
DB 121 FPDQNTYTPGPGVRYPLTFGWCYKLVPEPDKVEEANKGENTSLHHPVSLHGMDDPEREV 180

QY 181 LEWRFDLSRLAFHHVARELHPEYFKNC 206  
DB 181 LEWRFDLSRLAFHHVARELHPEYFKNC 206

RESULT 5

Q9PXW9 PRELIMINARY; PRT; 206 AA.

ID Q9PXW9  
AC Q9PXW9  
DT 01-MAY-2000 (TREMELrel. 13, Created)  
DT 01-MAY-2000 (TREMELrel. 13, Last sequence update)  
DT 01-DEC-2001 (TREMELrel. 19, Last annotation update)  
DE NEGATIVE FACTOR (F-PROTEIN) (27 KDA PROTEIN).  
GN NEF.  
OS Human immunodeficiency virus type 1.  
OC Viruses; Retroviral viruses; Retroviridae; Lentivirus.  
OX NCBI\_TaxID=11676;  
RN [1]  
RC SEQUENCE FROM N.A.  
RA Visco Comandini U., Yun Z., Paganelli R., Orlandi P., Salotti A.,  
RA Johansson B., Vahine A., Sonnerborg A.;  
RT "HIV-1 nef mutations and clinical long-term non progression: a  
RT molecular epidemiology study";  
RL Submitted (JUN-1997) to the EMBL/GenBank/DBJ databases.  
CC -1- SIMILARITY: NEF HAS GTPASE, GTP-BINDING AND AUTOPHOSPHORYLATING  
CC ACTIVITIES, IT SEEMS TO DOWN-REGULATE THE CD4 (T4) ANTIGEN.  
DR EMBL; AF011494; AAD01472.1;  
DR HSSP; P03406; IEFN.  
DR InterPro; IPR001558; F-protein.  
DR Pfam; PF00469; F-protein; 1.  
KW AIDS; GTP-binding; Myristate.  
SQ SEQUENCE 206 AA; 23333 MW; 2EBF2A6A3ECAF5EA CRC64;

Query Match 65.5%; Score 1105; DB 15; Length 206;  
Best Local Similarity 97.1%; Pred. No. 1.3e-92;  
Matches 200; Conservative 2; Mismatches 4; Indels 0; Gaps 0;

QY 1 MGKWSKSSVVGWPTVRRMRRAEPAADGVGAASRDLEKKGGLHSGRRQDILDLWIYHTGY 60  
DB 1 MGKWSKSSVIGWPAVRRMRRAEPAADGVGAASRDLEKKGGLHSGRRQDILDLWIYHTGY 60

QY 61 QEEEEVGFPVTPQVPLRPMTYKAADVLSHFLKEKGLHSGRRQDILDLWIYHTGY 120  
DB 61 QEEEEVGFPVTPQVPLRPMTYKAADVLSHFLKEKGLHSGRRQDILDLWIYHTGY 120

QY 121 FPDQNTYTPGPGVRYPLTFGWCYKLVPEPDKVEEANKGENTSLHHPVSLHGMDDPEREV 180  
DB 121 FPDQNTYTPGPGVRYPLTFGWCYKLVPEPDKVEEANKGENTSLHHPVSLHGMDDPEREV 180

QY 181 LEWRFDLSRLAFHHVARELHPEYFKNC 206  
DB 181 LEWRFDLSRLAFHHVARELHPEYFKNC 206

RESULT 7

Q9WM30 PRELIMINARY; PRT; 206 AA.

ID Q9WM30  
AC Q9WM30  
DT 01-NOV-1999 (TREMELrel. 12, Created)  
DT 01-NOV-1999 (TREMELrel. 12, Last sequence update)  
DT 01-DEC-2001 (TREMELrel. 19, Last annotation update)  
DE NEGATIVE FACTOR (F-PROTEIN) (27 KDA PROTEIN).  
GN NEF.  
OS Human immunodeficiency virus type 1.  
OC Viruses; Retroviral viruses; Retroviridae; Lentivirus.  
OX NCBI\_TaxID=11676;  
RN [1]  
RC SEQUENCE FROM N.A.  
RA Visco Comandini U., Yun Z., Paganelli R., Orlandi P., Salotti A.,  
RA Johansson B., Vahine A., Sonnerborg A.;  
RT "HIV-1 nef mutations and clinical long-term non progression: a  
RT molecular epidemiology study";  
RL Submitted (JUN-1997) to the EMBL/GenBank/DBJ databases.  
CC -1- SIMILARITY: NEF HAS GTPASE, GTP-BINDING AND AUTOPHOSPHORYLATING  
CC ACTIVITIES, IT SEEMS TO DOWN-REGULATE THE CD4 (T4) ANTIGEN.

```
DR EMBL: AF011469; AAD01447.1; -
DR HSP: P03406; IEFN.
DR InterPro: IPR001558; F-protein.
DR Pfam: PF00469; F-protein; 1.
DR AIDS: GTP-binding; Myristate.
SQ SEQUENCE 206 AA; 23345 MW; 0ED69927C2E03BB6 CRC64;

Query Match 65.4%; Score 1104; DB 15; Length 206;
Best Local Similarity 97.1%; Pred. No. 1.6e-92;
Matches 200; Conservative 1; Mismatches 5; Indels 0; Gaps 0;

QY 1 MGCKWSSVVGWPTVRRMRRAEPAADGVGAASRDLEKHGAISSNTAATNAACAWLEA 60
   |||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
Db 1 MGCKWSSVVGWPTVRRMRRAEPAADGVGAASRDLEKHGAISSNTAATNAACAWLEA 60
   |||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||

QY 61 QEEVEGFPVTPQVPLRPMTYKAADVLSHFLKEKGLGGLIHSQRQDILDLYHTQGY 120
   |||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
Db 61 QEEVEGFPVTPQVPLRPMTYKAADVLSHFLKEKGLGGLIHSQRQDILDLYHTQGY 120
   |||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||

QY 121 FPDQNYTTPGPGVRYPLTFCWCYKLVPEPDKVEEANKGENTSLHPVSLHGMDPPER 180
   |||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
Db 121 FPDQNYTTPGPGVRYPLTFCWCYKLVPEPDKVEEANKGENTSLHPVSLHGMDPPER 180
   |||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||

QY 181 LEWRFSRLAFHHVARELHPEYFKNC 206
   |||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
Db 181 LEWRFSRLAFHHVARELHPEYFKNC 206
   |||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||

RESULT 8
ID Q9PXW8 PRELIMINARY; PRT: 206 AA.
AC Q9PXW8;
DT 01-MAY-2000 (TrEMBLrel. 13, Created)
DT 01-MAY-2000 (TrEMBLrel. 13, Last sequence update)
DE 01-DEC-2001 (TrEMBLrel. 19, Last annotation update)
DE NEGATIVE FACTOR (F-PROTEIN) (27 KDA PROTEIN).
GN NEF.
OS Human immunodeficiency virus type 1.
OC Viruses; Retroviral viruses; Retroviridae; Lentivirus.
OC NCBI_TaxID=11676;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=93152025; PubMed=1301062;
RA Kohleisen B., Neumann M., Herrmann R., Brack-Werner R., Krohn K.J.,
   Ovod V., Ranki A., Erle V.;
RT "Cellular localization of Nef expressed in persistently HIV-1-infected
   low-producer astrocytes";
RL AIDS 6:1427-1436(1992).
CC -1- SIMILARITY: NEF HAS GTPASE, GTP-BINDING AND AUTOPHOSPHORYLATING
   ACTIVITIES, IT SEEMS TO DOWN-REGULATE THE CD4(14) ANTIGEN.
DR HSP: P03406; IEFN.
DR InterPro: IPR001558; F-protein.
DR Pfam: PF00469; F-protein; 1.
KW AIDS; GTP-binding; Myristate.
SQ SEQUENCE 206 AA; 23439 MW; 8478CF194D39D5B9 CRC64;

Query Match 65.2%; Score 1101; DB 15; Length 206;
Best Local Similarity 96.1%; Pred. No. 3e-92;
Matches 198; Conservative 4; Mismatches 4; Indels 0; Gaps 0;

QY 1 MGCKWSSVVGWPTVRRMRRAEPAADGVGAASRDLEKHGAISSNTAATNAACAWLEA 60
   |||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
Db 1 MSGRKSSVIGWPAVRRMRRAEPAADGVGAASRDLEKHGAISSNTAATNAACAWLEA 60
   |||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||

QY 61 QEEVEGFPVTPQVPLRPMTYKAADVLSHFLKEKGLGGLIHSQRQDILDLYHTQGY 120
   |||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
Db 61 QEEVEGFPVTPQVPLRPMTYKAADVLSHFLKEKGLGGLIHSQRQDILDLYHTQGY 120
   |||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||

QY 121 FPDQNYTTPGPGVRYPLTFCWCYKLVPEPDKVEEANKGENTSLHPVSLHGMDPPER 180
   |||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
Db 121 FPDQNYTTPGPGVRYPLTFCWCYKLVPEPDKVEEANKGENTSLHPVSLHGMDPPER 180
   |||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
```

```
QY 181 LEWRFSRLAFHHVARELHPEYFKNC 206
   |||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
Db 181 LEWRFSRLAFHHVARELHPEYFKNC 206
   |||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||

RESULT 9
ID Q90179 PRELIMINARY; PRT: 206 AA.
AC Q90179;
DT 01-NOV-1998 (TrEMBLrel. 08, Created)
DT 01-NOV-1998 (TrEMBLrel. 08, Last sequence update)
DT 01-DEC-2001 (TrEMBLrel. 19, Last annotation update)
DE NEGATIVE FACTOR (F-PROTEIN) (27 KDA PROTEIN).
GN NEF.
OS Human immunodeficiency virus type 1.
OC Viruses; Retroviral viruses; Retroviridae; Lentivirus.
OC NCBI_TaxID=11676;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=95074930; PubMed=7983770;
RA Fang H., Pincus S.H.;
RT "Unique insertion sequence and pattern of CD4 expression in variants
   selected with immunotoxins from human immunodeficiency virus type 1-
   infected T cells";
RL J. Virol. 69:75-81(1995).
RN [2]
RP SEQUENCE FROM N.A.
RA Fang H., Pincus S.H.;
RT "Spontaneous activation of human immunodeficiency virus type 1 in an
   immunotoxin-resistant variant T cell line";
RL Submitted (JUN-1998) to the EMBL/GenBank/DBJ databases.
CC -1- SIMILARITY: NEF HAS GTPASE, GTP-BINDING AND AUTOPHOSPHORYLATING
   ACTIVITIES, IT SEEMS TO DOWN-REGULATE THE CD4(14) ANTIGEN.
DR EMBL: AF070521; AAC28453.1; -
DR HSP: P03406; IEFN.
DR InterPro: IPR001558; F-protein.
DR Pfam: PF00469; F-protein; 1.
KW AIDS; GTP-binding; Myristate.
SQ SEQUENCE 206 AA; 23334 MW; 3E2B9C4017FDC6A8 CRC64;

Query Match 64.9%; Score 1096; DB 15; Length 206;
Best Local Similarity 95.6%; Pred. No. 8.6e-92;
Matches 197; Conservative 4; Mismatches 5; Indels 0; Gaps 0;

QY 1 MGCKWSSVVGWPTVRRMRRAEPAADGVGAASRDLEKHGAISSNTAATNAACAWLEA 60
   |||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
Db 1 MGCKWSSVIGWPAVRRMRRAEPAADGVGAASRDLEKHGAISSNTAATNAACAWLEA 60
   |||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||

QY 61 QEEVEGFPVTPQVPLRPMTYKAADVLSHFLKEKGLGGLIHSQRQDILDLYHTQGY 120
   |||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
Db 61 QEEVEGFPVTPQVPLRPMTYKAADVLSHFLKEKGLGGLIHSQRQDILDLYHTQGY 120
   |||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||

QY 121 FPDQNYTTPGPGVRYPLTFCWCYKLVPEPDKVEEANKGENTSLHPVSLHGMDPPER 180
   |||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
Db 121 FPDQNYTTPGPGVRYPLTFCWCYKLVPEPDKVEEANKGENTSLHPVSLHGMDPPER 180
   |||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||

QY 181 LEWRFSRLAFHHVARELHPEYFKNC 206
   |||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
Db 181 LEWRFSRLAFHHVARELHPEYFKNC 206
   |||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||

RESULT 10
ID Q85588 PRELIMINARY; PRT: 206 AA.
AC Q85588;
DT 01-NOV-1996 (TrEMBLrel. 01, Created)
DT 01-NOV-1996 (TrEMBLrel. 01, Last sequence update)
DT 01-DEC-2001 (TrEMBLrel. 19, Last annotation update)
DE NEGATIVE FACTOR (F-PROTEIN) (27 KDA PROTEIN).
GN NEF.
OS Aids-associated retrovirus.
```

OC Viruses; Retroviral viruses; Retroviridae.

```

OX NCBI_TaxID=11966;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=86067228; PubMed=2999715;
RA Ratner L., Starcich B., Josephs S.F., Hahn B.H., Reddy E.P., Arya S.K.,
RA Livak K.J., Petteway S.R.Jr., Pearson M.L., Haseltine W.A., Arya S.K.,
RA Wong-staal F.;
RT "Polymorphism of the 3' open reading frame of the virus associated
RT with the acquired immune deficiency syndrome, human T-lymphotropic
RT virus type III.";
RL Nucleic Acids Res. 13:8219-8229(1985).
CC -1- SIMILARITY: NEF HAS GTPASE, GTP-BINDING AND AUTOPHOSPHORYLATING
CC ACTIVITIES, IT SEEMS TO DOWN-REGULATE THE CD4(T4) ANTIGEN.
DR EMBL; X03188; CAA26947.1; -.
DR HSSP; P03406; IEFN.
DR InterPro; IPR001558; F-protein.
DR Pfam; PF00469; F-protein; 1.
DR AIDS; GTP-binding; Myristate.
KW AIDS; GTP-binding; Myristate.
SQ SEQUENCE 206 AA; 23419 MW; 3B25EB332A479A46 CRC64;

Query Match 64.9%; Score 1096; DB 15; Length 206;
Best Local Similarity 96.6%; Pred. No. 8.6e-92;
Matches 199; Conservative 4; Mismatches 3; Indels 0; Gaps 0;

Qy 1 MGGKWSKSSVVGWPTVRRMRRAEPAADGGAASRDLEKKGGLIHSQRQDILDWIYHTQGY 120
Db 1 MGGKWSKSSVVGWPTVRRMRRAEPAADGGAASRDLEKKGGLIHSQRQDILDWIYHTQGY 120
Qy 61 QEEVEGFPVTPQVPLRPMTYKAAVDLSHFLKKEKGLIHSQRQDILDWIYHTQGY 120
Db 61 QEEVEGFPVTPQVPLRPMTYKAAVDLSHFLKKEKGLIHSQRQDILDWIYHTQGY 120
Qy 121 FPDQNTYTPGVRPLTFGWCYKLVPEPKVEANKGENTSLHPVSLHGMDDPEREV 180
Db 121 FPDQNTYTPGVRPLTFGWCYKLVPEPKVEANKGENTSLHPVSLHGMDDPEREV 180
Qy 181 LEWRFDLSLAFHHVARELHPEYFKNC 206
Db 181 LEWRFDLSLAFHHVARELHPEYFKNC 206

RESULT 11
Q78244 ID Q78244 PRELIMINARY; PRT; 206 AA.
AC Q78244;
DT 01-NOV-1996 (TrEMBLrel. 01. Created)
DT 01-NOV-1996 (TrEMBLrel. 01. Last sequence update)
DT 01-DEC-2001 (TrEMBLrel. 19. Last annotation update)
DE NEGATIVE FACTOR (F-PROTEIN) (27 KDA PROTEIN).
GN NEF.
OS Human immunodeficiency virus type 1.
OC Viruses; Retroviral viruses; Retroviridae; Lentivirus.
OX NCBI_TaxID=11676;
RN [1]
RP SEQUENCE FROM N.A.
RA Carlini F., Federici M., Equestre M., Ricci S., Ratti G., Zibai Q.,
RA Verani P., Rossi G.B.;
RT "Sequence analysis of HIV-1 proviral DNA from a non producer
RT chronically infected HUT-78 cellular clone.";
RL J. Viral Diseases 1:40-55(1992).
RN [2]
RP SEQUENCE FROM N.A.
RX MEDLINE=89352106; PubMed=2765297;
RA Federico M., Titti F., Butto S., Orecchia A., Carlini F., Taddeo B.,
RA Macchi B., Mangiano N., Verani P., Rossi G.;
RT "Biological and molecular characterization of producer and non
RT producer clones from HUT-78 infected with a patient HIV isolate.";
RL AIDS Res. Hum. Retroviruses 5:385-396(1989).
RN [3]
RP SEQUENCE FROM N.A.
RA Titti F., Federico M., Butto S., Orecchia A., Carlini F., Taddeo B.,

```

```

RA Borsetti A., Saggio I., Verani P., Rossi G.;
RT "Variability of HIV-1 virus: characteristics of an infected but not
RT productive clone.";
RL Int. J. Immunopharmacol. 3:17-23(1990).
CC -1- SIMILARITY: NEF HAS GTPASE, GTP-BINDING AND AUTOPHOSPHORYLATING
CC ACTIVITIES, IT SEEMS TO DOWN-REGULATE THE CD4(T4) ANTIGEN.
DR EMBL; Z11530; CAA77629.1; -.
DR HSSP; P03406; IEFN.
DR InterPro; IPR001558; F-protein.
DR Pfam; PF00469; F-protein; 1.
DR AIDS; GTP-binding; Myristate.
KW AIDS; GTP-binding; Myristate.
SQ SEQUENCE 206 AA; 23398 MW; A31CC7244BBB2B43 CRC64;

Query Match 64.9%; Score 1096; DB 15; Length 206;
Best Local Similarity 96.6%; Pred. No. 8.6e-92;
Matches 199; Conservative 3; Mismatches 4; Indels 0; Gaps 0;

Qy 1 MGGKWSKSSVVGWPTVRRMRRAEPAADGGAASRDLEKKGGLIHSQRQDILDWIYHTQGY 120
Db 1 MGGKWSKSSVVGWPTVRRMRRAEPAADGGAASRDLEKKGGLIHSQRQDILDWIYHTQGY 120
Qy 61 QEEVEGFPVTPQVPLRPMTYKAAVDLSHFLKKEKGLIHSQRQDILDWIYHTQGY 120
Db 61 QEEVEGFPVTPQVPLRPMTYKAAVDLSHFLKKEKGLIHSQRQDILDWIYHTQGY 120
Qy 121 FPDQNTYTPGVRPLTFGWCYKLVPEPKVEANKGENTSLHPVSLHGMDDPEREV 180
Db 121 FPDQNTYTPGVRPLTFGWCYKLVPEPKVEANKGENTSLHPVSLHGMDDPEREV 180
Qy 181 LEWRFDLSLAFHHVARELHPEYFKNC 206
Db 181 LEWRFDLSLAFHHVARELHPEYFKNC 206

RESULT 12
Q9WLM7 ID Q9WLM7 PRELIMINARY; PRT; 206 AA.
AC Q9WLM7;
DT 01-NOV-1999 (TrEMBLrel. 12. Created)
DT 01-NOV-1999 (TrEMBLrel. 12. Last sequence update)
DT 01-DEC-2001 (TrEMBLrel. 19. Last annotation update)
DE NEGATIVE FACTOR (F-PROTEIN) (27 KDA PROTEIN).
GN NEF.
OS Human immunodeficiency virus type 1.
OC Viruses; Retroviral viruses; Retroviridae; Lentivirus.
OX NCBI_TaxID=11676;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=18-SW;
RA Visco Comandini U., Yun Z., Paganelli R., Orlandi P., Salotti A.,
RA Johansson B., Vahine A., Sonnerborg A.;
RT "HIV-1 nef mutations and clinical long-term non progression: a
RT molecular epidemiology study.";
RL Submitted (FEB-1998) to the EMBL/GenBank/DBJ databases.
CC -1- SIMILARITY: NEF HAS GTPASE, GTP-BINDING AND AUTOPHOSPHORYLATING
CC ACTIVITIES, IT SEEMS TO DOWN-REGULATE THE CD4(T4) ANTIGEN.
DR EMBL; AF047084; AAD02458.1; -.
DR HSSP; P03406; IEFN.
DR InterPro; IPR001558; F-protein.
DR Pfam; PF00469; F-protein; 1.
DR AIDS; GTP-binding; Myristate.
KW AIDS; GTP-binding; Myristate.
SQ SEQUENCE 206 AA; 23512 MW; E09E3BEF828A83C0 CRC64;

Query Match 64.7%; Score 1092; DB 15; Length 206;
Best Local Similarity 95.6%; Pred. No. 2e-91;
Matches 197; Conservative 7; Mismatches 2; Indels 0; Gaps 0;

Qy 1 MGGKWSKSSVVGWPTVRRMRRAEPAADGGAASRDLEKKGGLIHSQRQDILDWIYHTQGY 120
Db 1 MGGKWSKSSVVGWPTVRRMRRAEPAADGGAASRDLEKKGGLIHSQRQDILDWIYHTQGY 120

```

QY 61 QEEVEGFPVTPQVPLRPMTYKAAVDLSHFLKEKGLEGLIHSRRQDILDLDWIYHTQY 120  
DB 61 QEEVEGFPVTPQVPLRPMTYKAAVDLSHFLKEKGLEGLIHSRRQDILDLDWIYHTQY 120  
QY 121 FPDWQNTPGVRYPLTFGWCYKLVPEPKVEANKGENTSLHPVSLHGMDDPEREV 180  
DB 121 FPDWQNTPGVRYPLTFGWCYKLVPEPKVEANKGENTSLHPVSLHGMDDPEREV 180  
QY 181 LEWRDLSRLAFHHVARELHPEYFKNC 206  
DB 181 LEWRDLSRLAFHHVARELHPEYFKNC 206

RESULT 13  
Q74905 PRELIMINARY: PRT: 206 AA.  
AC Q74905;  
DT 01-NOV-1996 (TremBLrel. 01, Created)  
DT 01-NOV-1996 (TremBLrel. 01, Last sequence update)  
DT 01-DEC-2001 (TremBLrel. 19, Last annotation update)  
DE NEGATIVE FACTOR (F-PROTEIN) (27 KDA PROTEIN).  
GN NEF.  
OS Human immunodeficiency virus type 1.  
OC Viruses; Retroviral viruses; Retroviridae; Lentivirus.  
OX NCBI\_TaxID=11676;  
RN [1]  
RP SEQUENCE FROM N.A.  
RX MEDLINE=96400183; PubMed=8806559;  
RA Ratner L., Joseph T., Bandres J., Ghosh S., Heyden N.V., Templeton A.,  
Hahn B., Powderly W., Arens M.,  
RT "Sequence heterogeneity of Nef transcripts in HIV-1-infected subjects  
at different stages of disease."  
RL Virology 223:245-250(1996).  
CC -1- SIMILARITY: NEF HAS GTPASE, GTP-BINDING AND AUTOPHOSPHORYLATING  
ACTIVITIES, IT SEEMS TO DOWN-REGULATE THE CD4(T4) ANTIGEN.  
KW AIDS; GTP-binding; Myristate.  
SQ SEQUENCE 206 AA; 23617 MW; 5CAAB09C4730C38A CRC64;

Query Match 64.6%; Score 1090; DB 15; Length 206;  
Best Local Similarity 95.1%; Pred. No. 3e-91;  
Matches 196; Conservative 6; Mismatches 4; Indels 0; Gaps 0;

QY 1 MGGKWSKSSVGVPTVRMRRAEPAADGVGAASRDLEKHGAITSNTAATNACAWLEA 60  
DB 1 MGGKWSKSSVGVPTVRMRRAEPAADGVGAASRDLEKHGAITSNTAATNACAWLEA 60  
QY 61 QEEVEGFPVTPQVPLRPMTYKAAVDLSHFLKEKGLEGLIHSRRQDILDLDWIYHTQY 120  
DB 61 QEEVEGFPVTPQVPLRPMTYKAAVDLSHFLKEKGLEGLIHSRRQDILDLDWIYHTQY 120  
QY 121 FPDWQNTPGVRYPLTFGWCYKLVPEPKVEANKGENTSLHPVSLHGMDDPEREV 180  
DB 121 FPDWQNTPGVRYPLTFGWCYKLVPEPKVEANKGENTSLHPVSLHGMDDPEREV 180  
QY 181 LEWRDLSRLAFHHVARELHPEYFKNC 206  
DB 181 LEWRDLSRLAFHHVARELHPEYFKNC 206

RESULT 14  
Q89561 PRELIMINARY: PRT: 206 AA.  
AC Q89561;  
DT 01-NOV-1996 (TremBLrel. 01, Created)  
DT 01-NOV-1996 (TremBLrel. 01, Last sequence update)  
DT 01-DEC-2001 (TremBLrel. 19, Last annotation update)  
DE NEGATIVE FACTOR (F-PROTEIN) (27 KDA PROTEIN).  
GN NEF.

OS Human immunodeficiency virus type 1.  
OC Viruses; Retroviral viruses; Retroviridae; Lentivirus.  
OX NCBI\_TaxID=11676;  
RN [1]  
RP SEQUENCE FROM N.A.  
RX MEDLINE=96400183; PubMed=8806559;  
RA Ratner L., Joseph T., Bandres J., Ghosh S., Heyden N.V., Templeton A.,  
Hahn B., Powderly W., Arens M.,  
RT "Sequence heterogeneity of Nef transcripts in HIV-1-infected subjects  
at different stages of disease."  
RL Virology 223:245-250(1996).  
CC -1- SIMILARITY: NEF HAS GTPASE, GTP-BINDING AND AUTOPHOSPHORYLATING  
ACTIVITIES, IT SEEMS TO DOWN-REGULATE THE CD4(T4) ANTIGEN.  
KW AIDS; GTP-binding; Myristate.  
SQ SEQUENCE 206 AA; 23629 MW; BF467FB1B7147CDE CRC64;

Query Match 64.4%; Score 1087; DB 15; Length 206;  
Best Local Similarity 94.7%; Pred. No. 5.7e-91;  
Matches 195; Conservative 7; Mismatches 4; Indels 0; Gaps 0;

QY 1 MGGKWSKSSVGVPTVRMRRAEPAADGVGAASRDLEKHGAITSNTAATNACAWLEA 60  
DB 1 MGGKWSKSSVGVPTVRMRRAEPAADGVGAASRDLEKHGAITSNTAATNACAWLEA 60  
QY 61 QEEVEGFPVTPQVPLRPMTYKAAVDLSHFLKEKGLEGLIHSRRQDILDLDWIYHTQY 120  
DB 61 QEEVEGFPVTPQVPLRPMTYKAAVDLSHFLKEKGLEGLIHSRRQDILDLDWIYHTQY 120  
QY 121 FPDWQNTPGVRYPLTFGWCYKLVPEPKVEANKGENTSLHPVSLHGMDDPEREV 180  
DB 121 FPDWQNTPGVRYPLTFGWCYKLVPEPKVEANKGENTSLHPVSLHGMDDPEREV 180  
QY 181 LEWRDLSRLAFHHVARELHPEYFKNC 206  
DB 181 LEWRDLSRLAFHHVARELHPEYFKNC 206

RESULT 15  
Q74913 PRELIMINARY: PRT: 206 AA.  
AC Q74913;  
DT 01-NOV-1996 (TremBLrel. 01, Created)  
DT 01-NOV-1996 (TremBLrel. 01, Last sequence update)  
DT 01-DEC-2001 (TremBLrel. 19, Last annotation update)  
DE NEGATIVE FACTOR (F-PROTEIN) (27 KDA PROTEIN).  
GN NEF.  
OS Human immunodeficiency virus type 1.  
OC Viruses; Retroviral viruses; Retroviridae; Lentivirus.  
OX NCBI\_TaxID=11676;  
RN [1]  
RP SEQUENCE FROM N.A.  
RX MEDLINE=96400183; PubMed=8806559;  
RA Ratner L., Joseph T., Bandres J., Ghosh S., Heyden N.V., Templeton A.,  
Hahn B., Powderly W., Arens M.,  
RT "Sequence heterogeneity of Nef transcripts in HIV-1-infected subjects  
at different stages of disease."  
RL Virology 223:245-250(1996).  
CC -1- SIMILARITY: NEF HAS GTPASE, GTP-BINDING AND AUTOPHOSPHORYLATING  
ACTIVITIES, IT SEEMS TO DOWN-REGULATE THE CD4(T4) ANTIGEN.  
KW AIDS; GTP-binding; Myristate.  
SQ SEQUENCE 206 AA; 23649 MW; E5B97FA0B70FC175 CRC64;

Query Match 64.3%; Score 1085; DB 15; Length 206;  
Best Local Similarity 95.6%; Pred. No. 8.6e-91;  
Matches 196; Conservative 5; Mismatches 4; Indels 0; Gaps 0;

```
QY 1 MGGKSKSVVGVPTVRRMRRAEPAADGVGAASRDLEKHGAITSNTAATNAACAWLEA 60
DB 1 MGGKSKSVVGVPTVRRMRRAEPAADGVGAASRDLEKHGAITSNTAATNAACAWLEA 60
QY 61 QEEVEGFPVTPQVPLRPMPTKAAVDLSHFLKEKGLGGLIHSQRQDILDWYHTQGY 120
DB 61 QEEVEGFPVTPQVPLRPMPTKAAVDLSHFLKEKGLGGLIHSQRQDILDWYHTQGY 120
QY 121 FPDWNYTPGQVRYPLTFGWCYKLVPEPDKVEANKGENTSLHHPVSLHGMDDPEREV 180
DB 121 FPDWNYTPGQVRYPLTFGWCYKLVPEPDKVEANKGENTSLHHPVSLHGMDDPEREV 180
QY 181 LEWRDSDRLAFHHVARELHPEYFKN 205
DB 181 LEWRDSDRLAFHHVARELHPEYFKN 205

RESULT 16
Q74917 PRELIMINARY; PRT; 206 AA.
AC Q74917;
DT 01-NOV-1996 (TREMBlrel. 01, Created)
DT 01-NOV-1996 (TREMBlrel. 01, Last sequence update)
DT 01-DEC-2001 (TREMBlrel. 19, Last annotation update)
DE NEGATIVE FACTOR (F-PROTEIN) (27 KDA PROTEIN).
GN NEF.
OS Human immunodeficiency virus type 1.
OC Viruses; Retroid viruses; Retroviridae; Lentivirus.
OX NCBI_TaxID=11676;
RN [1]
RX MEDLINE=96400183; PubMed=8806559;
RA Hahn B., Joseph T., Bandres J., Ghosh S., Heyden N.V., Templeton A.,
RA Hahn B., Powderly W., Arens M.;
RT "Sequence heterogeneity of Nef transcripts in HIV-1-infected subjects
at different stages of disease.";
RL Virology 223:245-250(1996).
CC -!- SIMILARITY: NEF HAS GTPASE, GTP-BINDING AND AUTOPHOSPHORYLATING
ACTIVITIES, IT SEEMS TO DOWN-REGULATE THE CD4(T4) ANTIGEN.
CC
DR EMBL; U44454; AAB38206.1;
DR HSP; P03406; IEFN.
DR InterPro; IPR001558; F-protein.
DR Pfam; PF00469; F-protein; 1.
KW AIDS; GTP-binding; Myristate.
SQ SEQUENCE 206 AA; 23562 MW; 478BBFA675CB203 CRC64;
```

```
Query Match 64.0%; Score 1080; DB 15; Length 206;
Best Local Similarity 94.7%; Pred. No. 2.5e-90;
Matches 195; Conservative 5; Mismatches 6; Indels 0; Gaps 0;

QY 1 MGGKSKSVVGVPTVRRMRRAEPAADGVGAASRDLEKHGAITSNTAATNAACAWLEA 60
DB 1 MGGKSKSVVGVPTVRRMRRAEPAADGVGAASRDLEKHGAITSNTAATNAACAWLEA 60
QY 61 QEEVEGFPVTPQVPLRPMPTKAAVDLSHFLKEKGLGGLIHSQRQDILDWYHTQGY 120
DB 61 QEEVEGFPVTPQVPLRPMPTKAAVDLSHFLKEKGLGGLIHSQRQDILDWYHTQGY 120
QY 121 FPDWNYTPGQVRYPLTFGWCYKLVPEPDKVEANKGENTSLHHPVSLHGMDDPEREV 180
DB 121 FPDWNYTPGQVRYPLTFGWCYKLVPEPDKVEANKGENTSLHHPVSLHGMDDPEREV 180
QY 181 LEWRDSDRLAFHHVARELHPEYFKN 206
DB 181 LEWRDSDRLAFHHVARELHPEYFKN 206

RESULT 17
Q74914 PRELIMINARY; PRT; 206 AA.
AC Q74914;
DT 01-NOV-1996 (TREMBlrel. 01, Created)
```

```
DT 01-NOV-1996 (TREMBlrel. 01, Last sequence update)
DT 01-DEC-2001 (TREMBlrel. 19, Last annotation update)
DE NEGATIVE FACTOR (F-PROTEIN) (27 KDA PROTEIN).
GN NEF.
OS Human immunodeficiency virus type 1.
OC Viruses; Retroid viruses; Retroviridae; Lentivirus.
OX NCBI_TaxID=11676;
RN [1]
RX MEDLINE=96400183; PubMed=8806559;
RA Hahn B., Joseph T., Bandres J., Ghosh S., Heyden N.V., Templeton A.,
RA Hahn B., Powderly W., Arens M.;
RT "Sequence heterogeneity of Nef transcripts in HIV-1-infected subjects
at different stages of disease.";
RL Virology 223:245-250(1996).
CC -!- SIMILARITY: NEF HAS GTPASE, GTP-BINDING AND AUTOPHOSPHORYLATING
ACTIVITIES, IT SEEMS TO DOWN-REGULATE THE CD4(T4) ANTIGEN.
CC
DR EMBL; U44454; AAB38206.1;
DR HSP; P03406; IEFN.
DR InterPro; IPR001558; F-protein.
DR Pfam; PF00469; F-protein; 1.
KW AIDS; GTP-binding; Myristate.
SQ SEQUENCE 206 AA; 23615 MW; 08D97FA0B70FC17D CRC64;

Query Match 63.9%; Score 1079; DB 15; Length 206;
Best Local Similarity 95.1%; Pred. No. 3e-90;
Matches 195; Conservative 5; Mismatches 5; Indels 0; Gaps 0;

QY 1 MGGKSKSVVGVPTVRRMRRAEPAADGVGAASRDLEKHGAITSNTAATNAACAWLEA 60
DB 1 MGGKSKSVVGVPTVRRMRRAEPAADGVGAASRDLEKHGAITSNTAATNAACAWLEA 60
QY 61 QEEVEGFPVTPQVPLRPMPTKAAVDLSHFLKEKGLGGLIHSQRQDILDWYHTQGY 120
DB 61 QEEVEGFPVTPQVPLRPMPTKAAVDLSHFLKEKGLGGLIHSQRQDILDWYHTQGY 120
QY 121 FPDWNYTPGQVRYPLTFGWCYKLVPEPDKVEANKGENTSLHHPVSLHGMDDPEREV 180
DB 121 FPDWNYTPGQVRYPLTFGWCYKLVPEPDKVEANKGENTSLHHPVSLHGMDDPEREV 180
QY 181 LEWRDSDRLAFHHVARELHPEYFKN 205
DB 181 LEWRDSDRLAFHHVARELHPEYFKN 205

RESULT 18
Q74915 PRELIMINARY; PRT; 206 AA.
AC Q74915;
DT 01-NOV-1996 (TREMBlrel. 01, Created)
DT 01-NOV-1996 (TREMBlrel. 01, Last sequence update)
DT 01-DEC-2001 (TREMBlrel. 19, Last annotation update)
DE NEGATIVE FACTOR (F-PROTEIN) (27 KDA PROTEIN).
GN NEF.
OS Human immunodeficiency virus type 1.
OC Viruses; Retroid viruses; Retroviridae; Lentivirus.
OX NCBI_TaxID=11676;
RN [1]
RX MEDLINE=96400183; PubMed=8806559;
RA Hahn B., Joseph T., Bandres J., Ghosh S., Heyden N.V., Templeton A.,
RA Hahn B., Powderly W., Arens M.;
RT "Sequence heterogeneity of Nef transcripts in HIV-1-infected subjects
at different stages of disease.";
RL Virology 223:245-250(1996).
CC -!- SIMILARITY: NEF HAS GTPASE, GTP-BINDING AND AUTOPHOSPHORYLATING
ACTIVITIES, IT SEEMS TO DOWN-REGULATE THE CD4(T4) ANTIGEN.
CC
DR EMBL; U44455; AAB38207.1;
DR HSP; P03406; IEFN.
DR InterPro; IPR001558; F-protein.
DR Pfam; PF00469; F-protein; 1.
KW AIDS; GTP-binding; Myristate.
```



RL AIDS 0.0-0.0(2000).  
CC -1- SIMILARITY: NEF HAS GTPASE, GTP-BINDING AND AUTOPHOSPHORYLATING  
CC ACTIVITIES, IT SEEMS TO DOWN-REGULATE THE CD4(T4) ANTIGEN.  
DR EMBL: AF203197; AAF25320.1; -  
DR HSSP: P03406; LEFN.  
DR InterPro: IPR001558; F-protein.  
DR Pfam: PF00469; F-protein; 1.  
KW AIDS: GTP-binding; Myristate.  
SQ SEQUENCE 206 AA; 23510 MW; 081DF3A12E5A7576 CRC64;

Query Match 63.5%; Score 1072; DB 15; Length 206;  
Best Local Similarity 94.7%; Pred. No. 1.3e-89;  
Matches 195; Conservative 2; Mismatches 9; Indels 0; Gaps 0;

QY 1 MGKWSKSSVVGWPTVRRMRRAEPAADGVGAASRDLEKHGKAITSSNTAATNAACAWLEA 60  
DB 1 MGKWSKRIKDGWPAVRRMRRAEPAADGVGAASRDLEKHGKAITSSNTAATNAACAWLEA 60

QY 61 QEEEEVGFPVTPQVPLRPMTYKAADVLSHFLKKGLEGLIHSQRQDILDLYHTQY 120  
DB 61 QEEEEVGFPVTPQVPLRPMTYKAADVLSHFLKKGLEGLIHSQRQDILDLYHTQY 120

QY 121 FPDQNTYTPGQVRYPLTFGWCYKLVPEPDKVEEANKGENTSLHPVSLHGMDPDEREV 180  
DB 121 FPDQNTYTPGQVRYPLTFGWCYKLVPEPDKVEEANKGENTSLHPVSLHGMDPDEREV 180

QY 181 LEWRFDLSRLAFHHVARELHPEYFKNC 206  
DB 181 LEWRFDLSRLAFHHVARELHPEYFKNC 206

RESULT 22  
QY0595  
ID QY0595 PRELIMINARY; PRT: 206 AA.  
AC QY0595;  
DT 01-MAY-2000 (TrEMBLrel. 13, Created)  
DT 01-MAY-2000 (TrEMBLrel. 13, Last sequence update)  
DT 01-DEC-2001 (TrEMBLrel. 19, Last annotation update)  
DE NEGATIVE FACTOR (F-PROTEIN) (27 KDA PROTEIN).  
GN NEF.  
OS Human immunodeficiency virus type 1.  
OC Viruses; Retroviral viruses; Retroviridae; Lentivirus.  
OX NCBI\_TaxID=11676;  
RN [1]  
RP SEQUENCE FROM N.A.  
RC STRAIN-8-IT;  
RA Visco Comandini U., Yun Z., Paganelli R., Orlandi P., Salotti A.,  
RT Johansson B., Vahine A., Sonnerborg A.;  
RT "HIV-1 nef mutations and clinical long-term non progression: a  
RT molecular epidemiology study";  
RL Submitted (JUN-1997) to the EMBL/GenBank/DBJ databases.  
CC -1- SIMILARITY: NEF HAS GTPASE, GTP-BINDING AND AUTOPHOSPHORYLATING  
CC ACTIVITIES, IT SEEMS TO DOWN-REGULATE THE CD4(T4) ANTIGEN.  
DR EMBL: AF011470; AAD01448.1;  
DR HSSP: P03406; LEFN.  
DR InterPro: IPR001558; F-protein.  
DR Pfam: PF00469; F-protein; 1.  
KW AIDS: GTP-binding; Myristate.  
SQ SEQUENCE 206 AA; 23681 MW; 1C891C3AD230C6A8 CRC64;

Query Match 63.4%; Score 1071; DB 15; Length 208;  
Best Local Similarity 94.2%; Pred. No. 1.6e-89;  
Matches 196; Conservative 3; Mismatches 7; Indels 2; Gaps 1;

QY 1 MGKWSKSSVVGWPTVRRMRRAEPA--ADGVGAASRDLEKHGKAITSSNTAATNAACAWL 58  
DB 1 MGKWSKRSVIGWPAVRRMRRAEPAARXGAVSRDLEKHGKAITSSNTAATNAACAWL 60

QY 59 EAQEEVGFPVTPQVPLRPMTYKAADVLSHFLKKGLEGLIHSQRQDILDLYHTQ 118  
DB 61 EAQEEVGFPVTPQVPLRPMTYKAADVLSHFLKKGLEGLIHSQRQDILDLYHTQ 120

QY 119 GYFPDQNTYTPGQVRYPLTFGWCYKLVPEPDKVEEANKGENTSLHPVSLHGMDPDER 178  
DB 121 GYFPDQNTYTPGQVRYPLTFGWCYKLVPEPDKVEEANKGENTSLHPVSLHGMDPDER 180

QY 179 EVLEWRFDLSRLAFHHVARELHPEYFKNC 206  
DB 181 KVLEWRFDLSRLAFHHVARELHPEYFKNC 208

Query Match 63.5%; Score 1072; DB 15; Length 206;  
Best Local Similarity 93.7%; Pred. No. 1.3e-89;  
Matches 193; Conservative 5; Mismatches 8; Indels 0; Gaps 0;

QY 1 MGKWSKSSVVGWPTVRRMRRAEPAADGVGAASRDLEKHGKAITSSNTAATNAACAWLEA 60  
DB 1 MGKWSKRSVIGWPAVRRMRRAEPAADGVGAASRDLEKHGKAITSSNTAATNAACAWLEA 60

QY 61 QEEEEVGFPVTPQVPLRPMTYKAADVLSHFLKKGLEGLIHSQRQDILDLYHTQY 120  
DB 61 QEEEEVGFPVTPQVPLRPMTYKAADVLSHFLKKGLEGLIHSQRQDILDLYHTQY 120





```
ID O93010 PRELIMINARY; PRT; 206 AA.
AC O93010:
DT 01-NOV-1998 (TReMBLrel. 08, Created)
DT 01-NOV-1998 (TReMBLrel. 08, Last sequence update)
DT 01-DEC-2001 (TReMBLrel. 19, Last annotation update)
DE NEGATIVE FACTOR (F-PROTEIN) (27 KDA PROTEIN).
GN NEF.
OS Human immunodeficiency virus type 1.
OC Viruses; Retroid viruses; Retroviridae; Lentivirus.
OX NCBI_TaxID=11676;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN-PATIENT 9;
RX MEDLINE-98097260; PubMed-9436760;
RA Kang M.R., Cho Y.K., Chun J., Kim Y.B., Lee I., Lee H.J., Kim S.H.,
RA Kim Y.K., Yoon K., Yang J.M., Kim J.M., Shin Y.O., Kang C., Lee J.S.,
RA Choi K.W., Kim D.G., Fitch W.M., Kim S.;
RT "Phylogenetic analysis of the nef gene reveals a distinctive
RT monophyletic clade in Korean HIV-1 cases."
RL J. Acquir. Immune Defic. Syndr. Hum. Retrovirol. 17:58-68(1998).
CC -!- SIMILARITY: NEF HAS GTPASE, GTP-BINDING AND AUTOPHOSPHORYLATING
CC ACTIVITIES, IT SEEMS TO DOWN-REGULATE THE CD4(T4) ANTIGEN.
DR EMBL; AF063922; AAC17893.1; -
DR HSSP; P03406; 1EPN.
DR InterPro; IPR001558; F-protein.
DR Pfam; PF00469; F-protein; 1.
DR AIDS; GTP-binding; Myristate.
KW AIDS; GTP-binding; Myristate.
SQ SEQUENCE 206 AA; 23348 MW; 4D5AEDF55FAE93E3 CRC64;

Query Match 62.9%; Score 1062; DB 15; Length 206;
Best Local Similarity 94.6%; Pred. No. 1.1e-88;
Matches 194; Conservative 3; Mismatches 8; Indels 0; Gaps 0;

QY 1 MGGKWSKSVVGVPTVRERMRRAEPAADGVGAASRDLEKHGAITSNTAATNAACAWLEA 60
Db 1 MGGKWSKSVVGVPTVRERMRRAEPAADGVGAASRDLEKHGAITSNTAATNAACAWLEA 60
QY 61 QEEVEGVPVTPQVPLRPMTYKAAVDLSHFLKEKGLEGLIHSQRQDILDWIYHTQY 120
Db 61 QEEVEGVPVTPQVPLRPMTYKAAVDLSHFLKEKGLEGLIHSQRQDILDWIYHTQY 120
QY 121 FPDWQNTPGVRYPLTFGWCYKLVPEPDKVEANKGENTSLHPVSLHGMDDPEREV 180
Db 121 FPDWQNTPGVRYPLTFGWCYKLVPEPDKVEANKGENTSLHPVSLHGMDDPEREV 180
QY 181 LEWRDSDRLAFHHVARELHPEYFKN 205
Db 181 LEWRDSDRLAFHHVARELHPEYFKN 205
QY 181 LEWRDSDRLAFHHVARELHPEYKDC 206
Db 181 LEWRDSDRLAFHHVARELHPEYKDC 206

RESULT 28
Q93012 PRELIMINARY; PRT; 206 AA.
AC O93012:
DT 01-NOV-1998 (TReMBLrel. 08, Created)
DT 01-NOV-1998 (TReMBLrel. 08, Last sequence update)
DT 01-DEC-2001 (TReMBLrel. 19, Last annotation update)
DE NEGATIVE FACTOR (F-PROTEIN) (27 KDA PROTEIN).
GN NEF.
OS Human immunodeficiency virus type 1.
OC Viruses; Retroid viruses; Retroviridae; Lentivirus.
OX NCBI_TaxID=11676;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN-PATIENT 27;
RA McPhee D.A., Greenway A.L., Holloway G., Smith K., Deacon N.,
RA Pemberton L., Brew B.J.;
RT "Anomalies in nef expression within the central nervous system of HIV-
RT 1 positive individuals/AIDS patients with or without AIDS dementia
RT complex."
RL J. Neurovirol. 4:0-0(1998).
CC -!- SIMILARITY: NEF HAS GTPASE, GTP-BINDING AND AUTOPHOSPHORYLATING
```

```
CC ACTIVITIES, IT SEEMS TO DOWN-REGULATE THE CD4(T4) ANTIGEN.
DR EMBL; AF064673; AAC18374.1; -
DR HSSP; P03406; 1EPN.
DR InterPro; IPR001558; F-protein.
DR Pfam; PF00469; F-protein; 1.
DR AIDS; GTP-binding; Myristate.
KW AIDS; GTP-binding; Myristate.
SQ SEQUENCE 206 AA; 23472 MW; 2255A447ECE85456 CRC64;

Query Match 62.9%; Score 1061; DB 15; Length 206;
Best Local Similarity 91.7%; Pred. No. 1.3e-88;
Matches 189; Conservative 12; Mismatches 5; Indels 0; Gaps 0;

QY 1 MGGKWSKSVVGVPTVRERMRRAEPAADGVGAASRDLEKHGAITSNTAATNAACAWLEA 60
Db 1 MGGKWSKSVVGVPTVRERMRRAEPAADGVGAASRDLEKHGAITSNTAATNAACAWLEA 60
QY 61 QEEVEGVPVTPQVPLRPMTYKAAVDLSHFLKEKGLEGLIHSQRQDILDWIYHTQY 120
Db 61 QEEVEGVPVTPQVPLRPMTYKAAVDLSHFLKEKGLEGLIHSQRQDILDWIYHTQY 120
QY 121 FPDWQNTPGVRYPLTFGWCYKLVPEPDKVEANKGENTSLHPVSLHGMDDPEREV 180
Db 121 FPDWQNTPGVRYPLTFGWCYKLVPEPDKVEANKGENTSLHPVSLHGMDDPEREV 180
QY 181 LEWRDSDRLAFHHVARELHPEYFKN 206
Db 181 LEWRDSDRLAFHHVARELHPEYKDC 206

RESULT 29
Q90597 PRELIMINARY; PRT; 206 AA.
AC Q90597:
DT 01-MAY-2000 (TReMBLrel. 13, Created)
DT 01-MAY-2000 (TReMBLrel. 13, Last sequence update)
DT 01-DEC-2001 (TReMBLrel. 19, Last annotation update)
DE NEGATIVE FACTOR (F-PROTEIN) (27 KDA PROTEIN).
GN NEF.
OS Human immunodeficiency virus type 1.
OC Viruses; Retroid viruses; Retroviridae; Lentivirus.
OX NCBI_TaxID=11676;
RN [1]
RP SEQUENCE FROM N.A.
RA Wang B., Sakena N.K.;
RT "HIV-1 Strains from a cohort of American subjects reveal the presence
RT of a V2 region extension unique to slow progressors; and non-
RT progressors."
RL AIDS 0:0-0(2000).
CC -!- SIMILARITY: NEF HAS GTPASE, GTP-BINDING AND AUTOPHOSPHORYLATING
CC ACTIVITIES, IT SEEMS TO DOWN-REGULATE THE CD4(T4) ANTIGEN.
DR EMBL; AF203196; AAF25318.1; -
DR HSSP; P03406; 1EPN.
DR InterPro; IPR001558; F-protein.
DR Pfam; PF00469; F-protein; 1.
DR AIDS; GTP-binding; Myristate.
KW AIDS; GTP-binding; Myristate.
SQ SEQUENCE 206 AA; 23599 MW; 9C887A58541A9905 CRC64;

Query Match 62.8%; Score 1060; DB 15; Length 206;
Best Local Similarity 92.7%; Pred. No. 1.6e-88;
Matches 191; Conservative 7; Mismatches 8; Indels 0; Gaps 0;

QY 1 MGGKWSKSVVGVPTVRERMRRAEPAADGVGAASRDLEKHGAITSNTAATNAACAWLEA 60
Db 1 MGGKWSKSVVGVPTVRERMRRAEPAADGVGAASRDLEKHGAITSNTAATNAACAWLEA 60
QY 61 QEEVEGVPVTPQVPLRPMTYKAAVDLSHFLKEKGLEGLIHSQRQDILDWIYHTQY 120
Db 61 QEEVEGVPVTPQVPLRPMTYKAAVDLSHFLKEKGLEGLIHSQRQDILDWIYHTQY 120
QY 121 FPDWQNTPGVRYPLTFGWCYKLVPEPDKVEANKGENTSLHPVSLHGMDDPEREV 180
Db 121 FPDWQNTPGVRYPLTFGWCYKLVPEPDKVEANKGENTSLHPVSLHGMDDPEREV 180
```



GenCore version 4.5  
Copyright (c) 1993 - 2000 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: August 26, 2002, 08:12:25 ; Search time 99.16 Seconds  
(without alignments)  
460.380 Million cell updates/sec

Title: US-09-509-239-17  
Perfect score: 2255  
Sequence: 1 CSHSSNMANTQMSKDKIII.....QSRGDPGPKETSGHHHHH 411

Scoring table: BLOSUM62  
Gapop 10.0 , Gapext 0.5

Searched: 747574 seqs, 111073796 residues

Total number of hits satisfying chosen parameters: 747574

Minimum DB seq length: 0  
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

Database : A\_Geneseq\_032802.\*  
1: /SIDS1/gcgdata/hold-geneseq/geneseq-emb1/AA1980.DAT.\*  
2: /SIDS1/gcgdata/hold-geneseq/geneseq-emb1/AA1981.DAT.\*  
3: /SIDS1/gcgdata/hold-geneseq/geneseq-emb1/AA1982.DAT.\*  
4: /SIDS1/gcgdata/hold-geneseq/geneseq-emb1/AA1983.DAT.\*  
5: /SIDS1/gcgdata/hold-geneseq/geneseq-emb1/AA1984.DAT.\*  
6: /SIDS1/gcgdata/hold-geneseq/geneseq-emb1/AA1985.DAT.\*  
7: /SIDS1/gcgdata/hold-geneseq/geneseq-emb1/AA1986.DAT.\*  
8: /SIDS1/gcgdata/hold-geneseq/geneseq-emb1/AA1987.DAT.\*  
9: /SIDS1/gcgdata/hold-geneseq/geneseq-emb1/AA1988.DAT.\*  
10: /SIDS1/gcgdata/hold-geneseq/geneseq-emb1/AA1989.DAT.\*  
11: /SIDS1/gcgdata/hold-geneseq/geneseq-emb1/AA1990.DAT.\*  
12: /SIDS1/gcgdata/hold-geneseq/geneseq-emb1/AA1991.DAT.\*  
13: /SIDS1/gcgdata/hold-geneseq/geneseq-emb1/AA1992.DAT.\*  
14: /SIDS1/gcgdata/hold-geneseq/geneseq-emb1/AA1993.DAT.\*  
15: /SIDS1/gcgdata/hold-geneseq/geneseq-emb1/AA1994.DAT.\*  
16: /SIDS1/gcgdata/hold-geneseq/geneseq-emb1/AA1995.DAT.\*  
17: /SIDS1/gcgdata/hold-geneseq/geneseq-emb1/AA1996.DAT.\*  
18: /SIDS1/gcgdata/hold-geneseq/geneseq-emb1/AA1997.DAT.\*  
19: /SIDS1/gcgdata/hold-geneseq/geneseq-emb1/AA1998.DAT.\*  
20: /SIDS1/gcgdata/hold-geneseq/geneseq-emb1/AA1999.DAT.\*  
21: /SIDS1/gcgdata/hold-geneseq/geneseq-emb1/AA2000.DAT.\*  
22: /SIDS1/gcgdata/hold-geneseq/geneseq-emb1/AA2001.DAT.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

#### SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	2255	100.0	411	22 AAG63235	Amino acid sequenc
2	2246	99.6	413	20 AAY02355	A representative L
3	2246	99.6	413	22 AAG63237	Amino acid sequenc
4	2242	99.4	411	20 AAY02353	A representative L
5	1706.5	75.7	324	20 AAY02352	A representative L
6	1706.5	75.7	324	22 AAG63234	Amino acid sequenc
7	1697.5	75.3	326	20 AAY02354	A representative L
8	1697.5	75.3	326	22 AAG63236	Amino acid sequenc
9	1688	74.9	302	20 AAY02351	A representative H
10	1688	74.9	302	22 AAG63233	Amino acid sequenc
11	1675	74.3	302	20 AAY02357	A representative H

12	1675	74.3	302	22 AAG63239	Amino acid sequenc
13	1139.5	50.5	215	20 AAY02349	A representative H
14	1139.5	50.5	215	22 AAG63231	Amino acid sequenc
15	1120	49.7	206	21 AAY50795	Human NEF protein/
16	1116	49.5	206	14 AAR38893	Nef protein of HIV
17	1115	49.4	206	21 AAB10054	HIV-1 nef protein.
18	1114	49.4	206	20 AAW93326	HIV-1 nef protein
19	1111	49.3	206	7 AAP61515	Sequence of E' pro
20	1107	49.1	216	7 AAP60423	Sequence of LAV v1
21	1092	48.4	206	20 AAW90179	HTLV-III E' protei
22	1011	44.8	216	22 AAE04960	HIV-1 Jrf1 Nef pro
23	997	44.2	210	19 AAW52113	Protein 6 containe
24	997	44.2	210	21 AAY77299	HIV-1 Jrf1 Nef (G2
25	996	44.2	217	22 AAE04962	HIV-1 Jrf1 Nef OVI
26	983	43.6	210	12 AAR12262	Human tPA leader p
27	978	43.4	237	22 AAE04961	Human tPA leader p
28	968	42.9	237	22 AAE04963	HIV-1 non-subtype
29	952	42.2	206	21 AAB69363	HIV-1 non-subtype
30	948.5	42.1	3025	22 AAB86169	HIV-1 non-subtype
31	934	41.4	206	21 AAB69361	Sequence encoded b
32	927.5	41.1	206	9 AAP81859	HIV-1 non-subtype
33	910.5	40.4	207	21 AAB69359	Sequence deduced f
34	903.5	40.1	207	11 AAR08407	HIV-1 non-subtype
35	901.5	40.0	207	21 AAB69364	HIV-1 non-subtype
36	891.5	39.7	208	21 AAB69356	HIV-1 non-subtype
37	893.5	39.6	219	21 AAB69362	HIV-1 non-subtype
38	893.5	39.6	217	21 AAB69365	HIV-1 non-subtype
39	885.5	39.3	207	21 AAB69366	HIV-1 non-subtype
40	878.5	39.0	207	21 AAB69357	HIV-1 non-subtype
41	867	38.4	206	21 AAB69358	HIV-1 non-subtype
42	860	38.1	206	21 AAB69360	HIV isolate LAV.MA
43	834.5	37.0	209	19 AAW72998	Sequence encoded b
44	831.5	36.9	209	9 AAP81866	HIV-1 strain YBF30
45	796	35.3	212	19 AAW69481	

#### ALIGNMENTS

RESULT 1  
AAG63235  
ID AAG63235 standard; Protein; 411 AA.  
AC AAG63235;  
XX  
XX 01-OCT-2001 (first entry)  
DT  
XX Amino acid sequence of a His tagged LipD-Nef-Tat linked protein.  
XX HIV; nef gene; tat gene; vaccine; gp120 gene; HIV viral load.  
XX  
XX Synthetic.  
OS Human immunodeficiency virus.  
XX  
XX Key Location/Qualifiers  
FT Peptide 1..109  
FT /note= "protD fusion partner"  
XX  
XX WO200154719-A2.  
XX  
XX 02-AUG-2001.  
XX  
XX 29-JAN-2001; 2001WO-EP00944.  
XX  
XX 31-JAN-2000; 2000GB-0002200.  
PR 14-APR-2000; 2000GB-0009336.  
PR 06-JUN-2000; 2000GB-0013806.  
PR 28-JUN-2000; 2000WO-EP05998.  
XX  
XX (SMIK ) SMITHKLINE BEECHAM BIOLOGICALS.  
XX  
XX Voss G;  
XX

DR WPI: 2001-476172/51.  
 DR N-PSDB; AAH42879.  
 XX  
 PT New use of a human immunodeficiency virus (HIV) Tat, or Nef, or Nef, or Nef  
 PT linked to Tat (Nef-Tat) protein or polynucleotide and an HIV gp120  
 PT protein or polynucleotide for the manufacture of a vaccine -  
 XX  
 PS Disclosure; Fig 1; 90pp; English.  
 XX  
 CC The present sequence represents a His-tagged Nef-Tat linked protein of  
 CC HIV, with a lipidation signal sequence (LipoD) which is removed after  
 CC processing and a ProD fusion partner. The protein is expressed in  
 CC Escherichia coli, and is used to produce the vaccine of the invention.  
 CC The specification describes the use of HIV Tat, HIV Nef, or Nef-Tat;  
 CC and HIV gp120 in the manufacture of a vaccine. The vaccine is used for  
 CC the prophylactic or therapeutic immunization of humans against HIV.  
 CC Tat, Nef or Nef-Tat act in synergy with gp120 in the treatment and  
 CC prevention of HIV. The vaccine reduces the HIV viral load in HIV  
 CC infected humans and results in a maintenance of CD4+ levels over those  
 CC levels found in the absence of vaccination with HIV Tat, Nef or Nef-Tat  
 CC and HIV gp120.  
 XX  
 SQ Sequence 411 AA;  
 Query Match 100.0%; Score 2255; DB 22; Length 411;  
 Best Local Similarity 100.0%; Pred. No. 1.7e-205;  
 Matches 411; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
 QY 1 CSHSSNMANTQMSKDIHHRGASGYLPEHTLESKALAFQAQADYLEQDLAMTKDGR 60  
 DB 1 CSHSSNMANTQMSKDIHHRGASGYLPEHTLESKALAFQAQADYLEQDLAMTKDGR 60  
 QY 61 VVHDFHDLGLTDVAKKPPHRRKDGYYVIDFTLKEIQSLEMTENFETMGKWSKSSV 120  
 DB 61 VVHDFHDLGLTDVAKKPPHRRKDGYYVIDFTLKEIQSLEMTENFETMGKWSKSSV 120  
 QY 121 GWPTVRMRRAEPAADGVGAASRDLEKHGAITSSNTAATNAACAWLEAQEEVEGFPVT 180  
 DB 121 GWPTVRMRRAEPAADGVGAASRDLEKHGAITSSNTAATNAACAWLEAQEEVEGFPVT 180  
 QY 181 PVPURPMYKAAVDLSHFLKEKGLGLIHSQRQDILDWYHTQGYFPDQWNTTGP 240  
 DB 181 PVPURPMYKAAVDLSHFLKEKGLGLIHSQRQDILDWYHTQGYFPDQWNTTGP 240  
 QY 241 GVRYPITFGWCYKLVPEPDKVEEANKGENTSLHPVSLHGMDDPEREVLEWRFDSRLAF 300  
 DB 241 GVRYPITFGWCYKLVPEPDKVEEANKGENTSLHPVSLHGMDDPEREVLEWRFDSRLAF 300  
 QY 301 HHVARELHPEYFNKCTSEPVDPRLPEKHPGSGQPKTACTNCCYKCCFHCQVCFTKALG 360  
 DB 301 HHVARELHPEYFNKCTSEPVDPRLPEKHPGSGQPKTACTNCCYKCCFHCQVCFTKALG 360  
 QY 361 ISYGRKKRRQRPRPGSQTHQVSLSKQPTSSQSRGDPGPKETSGHHHHH 411  
 DB 361 ISYGRKKRRQRPRPGSQTHQVSLSKQPTSSQSRGDPGPKETSGHHHHH 411  
 RESULT 2  
 ID AAY02355 standard; Protein; 413 AA.  
 AC AAY02355;  
 XX  
 XX 09-JUL-1999 (first entry)  
 DT  
 XX A representative LipoD-Tat fusion protein.  
 DE  
 XX HIV nef gene; fusion protein; HIV nef gene; Nef protein; Tat protein;  
 KW vaccine; HIV infection; protein D.  
 XX  
 XX Synthetic.  
 OS Human immunodeficiency virus type 1.

XX WO9916884-A1.  
 XX  
 PD 08-APR-1999.  
 XX  
 PF 17-SEP-1998; 98WO-EP06040.  
 XX  
 PR 26-SEP-1997; 97GB-0020585.  
 XX  
 PA (SMIK ) SMITHKLINE BEECHAM BIOLOGICALS.  
 XX  
 PI Bruck C, Godart SAG, Marchand M;  
 XX  
 DR WPI: 1999-302282/25.  
 DR N-PSDB; AAX35691.  
 XX  
 PT HIV Tat or Nef protein linked to a fusion partner  
 PS Disclosure; Fig 2; 66pp; English.  
 XX  
 CC The present sequence represents a fusion protein comprising LipoD-HIV-1  
 CC Tat. The protein is exemplified the fusion proteins of  
 CC the invention. The specification also describes fusion proteins  
 CC comprising HIV-1 Nef protein. The fusion protein can be used in a  
 CC vaccine to prevent HIV infection.  
 XX  
 SQ Sequence 413 AA;  
 Query Match 99.6%; Score 2246; DB 20;  
 Best Local Similarity 100.0%; Pred. No. 1.3e-204;  
 Matches 410; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
 QY 2 SSSNSNMANTQMSKDIHHRGASGYLPEHTLESKALAFQAQADYLEQDLAMTKDGR 61  
 DB 4 SSSNSNMANTQMSKDIHHRGASGYLPEHTLESKALAFQAQADYLEQDLAMTKDGR 63  
 QY 62 VVHDFHDLGLTDVAKKPPHRRKDGYYVIDFTLKEIQSLEMTENFETMGKWSKSSV 121  
 DB 64 VVHDFHDLGLTDVAKKPPHRRKDGYYVIDFTLKEIQSLEMTENFETMGKWSKSSV 123  
 QY 122 WPTVRMRRAEPAADGVGAASRDLEKHGAITSSNTAATNAACAWLEAQEEVEGFPVT 181  
 DB 124 WPTVRMRRAEPAADGVGAASRDLEKHGAITSSNTAATNAACAWLEAQEEVEGFPVT 183  
 QY 182 QVPLPMTYKAAVDLSHFLKEKGLGLIHSQRQDILDWYHTQGYFPDQWNTTGP 241  
 DB 184 QVPLPMTYKAAVDLSHFLKEKGLGLIHSQRQDILDWYHTQGYFPDQWNTTGP 243  
 QY 242 VRYPLTFGWCYKLVPEPDKVEEANKGENTSLHPVSLHGMDDPEREVLEWRFDSRLAF 301  
 DB 244 VRYPLTFGWCYKLVPEPDKVEEANKGENTSLHPVSLHGMDDPEREVLEWRFDSRLAF 303  
 QY 302 HVARELHPEYFNKCTSEPVDPRLPEKHPGSGQPKTACTNCCYKCCFHCQVCFTKALG 361  
 DB 304 HVARELHPEYFNKCTSEPVDPRLPEKHPGSGQPKTACTNCCYKCCFHCQVCFTKALG 363  
 QY 362 SYGRKKRRQRPRPGSQTHQVSLSKQPTSSQSRGDPGPKETSGHHHHH 411  
 DB 364 SYGRKKRRQRPRPGSQTHQVSLSKQPTSSQSRGDPGPKETSGHHHHH 413  
 RESULT 3  
 ID AAG63237 standard; Protein; 413 AA.  
 XX  
 AC AAG63237;  
 XX  
 DT 01-OCT-2001 (first entry)  
 XX  
 DE Amino acid sequence of a His tagged ProD-Nef-Tat fusion protein.  
 XX HIV; nef gene; tat gene; vaccine; gp120 gene; HIV viral load.

XX Synthetic.  
OS Human immunodeficiency virus.  
XX Key Location/Qualifiers  
FH Peptide 1..111  
FT /note= "ProTid fusion partner"  
XX  
XX WO200154719-A2.  
XX  
XX 02-AUG-2001.  
XX  
XX 29-JAN-2001; 2001WO-EP00944.  
XX  
XX 31-JAN-2000; 2000GB-0002200.  
XX 14-APR-2000; 2000GB-0009336.  
XX 06-JUN-2000; 2000GB-0013806.  
XX 28-JUN-2000; 2000WO-EP05998.  
XX  
XX (SMIK ) SMITHKLINE BEECHAM BIOLOGICALS.  
XX  
XX Voss G;  
XX  
XX WPI; 2001-476172/51.  
XX N-PSDB; AAH42881.  
XX  
XX New use of a human immunodeficiency virus (HIV) Tat, or Nef, or Nef  
XX linked to Tat (Nef-Tat) protein or polynucleotide and an HIV gp120  
XX protein or polynucleotide for the manufacture of a vaccine -  
XX  
XX Disclosure; Fig 1; 90pp; English.  
XX  
XX The present sequence represents a His-tagged ProTid-Nef-Tat fusion  
XX protein. The protein is expressed in Escherichia coli, and is used to  
XX produce the vaccine of the invention. The specification describes  
XX the use of HIV Tat, HIV Nef, or Nef-Tat; and HIV gp120 in the  
XX manufacture of a vaccine. The vaccine is used for the prophylactic or  
XX therapeutic immunization of humans against HIV. Tat, Nef or Nef-Tat act  
XX in synergy with gp120 in the treatment and prevention of HIV. The  
XX vaccine reduces the HIV viral load in HIV infected humans and results  
XX in a maintenance of CD4+ levels over those levels found in the absence  
XX of vaccination with HIV Tat, Nef or Nef-Tat and HIV gp120.  
XX  
XX Sequence 413 AA;  
SQ

Query Match 99.6%; Score 2246; DB 22; Length 413;  
Best Local Similarity 100.0%; Pred. No. 1.3e-204;  
Matches 410; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 2 SSSSNMANTQMSDKIIIAHRCASGYLPEHTLESKALAFQAQADYLEQDLAMTKDGLV 61  
DB 4 ssssnmantqmsdkIIIAHRCASGYLPEHTLESKALAFQAQADYLEQDLAMTKDGLV 63  
QY 62 VIHDFLDGLTDVAKKFPFHRHRKDGYYVDFTLKEIQSLEMTENFETMGKWSKSSVVG 121  
DB 64 vindhfdlgltdvakkfphrhrkdgryyvdftlkeiqslemtentfmggkwsksvvg 123  
QY 122 WPTVRMRRAEPAADGVGAASRDLEKHGATSSNTAATNAACAWLEAQEEVEGFPVTP 181  
DB 124 wptvrermrraepaadvgaasrdlekghaigtssntaatnaacawleaqeeveegfpvtp 183  
QY 182 QVPLRPMTYKAAVDLSHFLKEKGLSLHSQRQDILDLWIYHTOGYFPDQWNYTPGPG 241  
DB 184 qvplrpmtykaavdlshflkexgglslhsqrqgdildlwihtdgyfpdwgnytpgpg 243  
QY 242 VRYPLTFGCYKLVPEPKVEANKGENTSLHPSVLSHGMDPPERLEWRFDSRLAPH 301  
DB 244 vrypltfgcylkvppepkveankgentslhpsvlsghmdpperlewrfdsrldph 303  
QY 302 HVARELHPYFNCTSEPDRLPEPKHFGSQPKTACTNCKYKCCCFHCQVCITKALGI 361  
DB 304 hvarelhpeyfnctsepdrlpepkhfgsqpkactnckykcccfhcqvcitkalg 363

QY 362 SYGRKKRRQRRRPOGSGTHQVSLSKQPTSQSRGDPGTGPKETSGHHHHH 411  
DB 364 sygrkkrrrrppgsgthqvsksqptsqsrqgdpgtgpketsghhhhh 413  
RESULT 4  
AA02353  
ID AAY02353 standard; Protein; 411 AA.  
XX  
XX AC AAY02353;  
XX  
XX DT 09-JUL-1999 (first entry)  
XX  
XX DE A representative LipoD-Tat-His fusion protein.  
XX  
XX KW HIV nef gene; fusion protein; HIV nef gene; Nef protein; Tat protein;  
XX KW vaccine; HIV infection; protein D.  
XX  
XX OS Synthetic.  
XX OS Human immunodeficiency virus type 1.  
XX  
XX PN WO9916884-A1.  
XX  
XX PD 08-APR-1999.  
XX  
XX PF 17-SEP-1998; 98WO-EP06040.  
XX  
XX PR 26-SEP-1997; 97GB-0020585.  
XX  
XX PA (SMIK ) SMITHKLINE BEECHAM BIOLOGICALS.  
XX  
XX PI Bruck C, Godart SAG, Marchand M;  
XX  
XX DR WPI; 1999-302282/25.  
XX  
XX DR N-PSDB; AAX35689.  
XX  
XX PT HIV Tat or Nef protein linked to a fusion partner  
XX  
XX PS Disclosure; Fig 2; 66pp; English.  
XX  
XX The present sequence represents a fusion protein comprising Lipod-HIV-1  
XX Tat-His. The protein is exemplified the fusion proteins of  
XX the invention. The specification also describes fusion proteins  
XX comprising HIV-1 Nef protein. The fusion protein can be used in a  
XX vaccine to prevent HIV infection.  
XX  
XX SQ Sequence 411 AA;

Query Match 99.4%; Score 2242; DB 20; Length 411;  
Best Local Similarity 99.8%; Pred. No. 3e-204;  
Matches 410; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 CSSHSSNMANTQMSDKIIIAHRCASGYLPEHTLESKALAFQAQADYLEQDLAMTKDGLR 60  
DB 1 csshssnmantqmsdkIIIAHRCASGYLPEHTLESKALAFQAQADYLEQDLAMTKDGLR 60  
QY 61 VIHDFLDGLTDVAKKFPFHRHRKDGYYVDFTLKEIQSLEMTENFETMGKWSKSSV 120  
DB 61 vvhdfldgltdvakkfphrhrkdgryyvdftlkeiqslemtentfmggkwsksv 120  
QY 121 GWPVTVRMRRAEPAADGVGAASRDLEKHGATSSNTAATNAACAWLEAQEEVEGFPVT 180  
DB 121 gwpvtvrermrraepaadvgaasrdlekghaigtssntaatnaacawleaqeevegfpvt 180  
QY 181 PQVPLRPMTYKAAVDLSHFLKEKGLSLHSQRQDILDLWIYHTOGYFPDQWNYTPGP 240  
DB 181 pqvplrpmtykaavdlshflkexgglslhsqrqgdildlwihtdgyfpdwgnytpgp 240  
QY 241 GVRYPPLTFGCYKLVPEPKVEANKGENTSLHPSVLSHGMDPPERLEWRFDSRLAF 300  
DB 241 gvryppltfgcylkvppepkveankgentslhpsvlsghmdpperlewrfdsrlaf 300

QY 301 HHVARELHPEYFKNCTSEPVDRLEPWKHGSPQKPTACTNCTCYCKKCCFHCQVCFITKALG 360  
 Db 301 hhvarelhpeyfkntsepvdrlepwkhgspqktactnctcyckkccfhcqvcfitkalg 360  
 QY 361 ISYGRKKRRRPPQSGTHQVSLSKOPTSOSRGDPTGPKETSGHHHHH 411  
 Db 361 isygrkkrrrrppqsgthqvskskptdsrgdptgpketsghhhhh 411

RESULT 5  
 AAY02352  
 ID AAY02352 standard; Protein; 324 AA.  
 XX  
 AC AAY02352;  
 DT 09-JUL-1999 (first entry)  
 XX  
 DE A representative LipD-Nef-His fusion protein.  
 XX  
 KW HIV nef gene; fusion protein; HIV nef gene; Nef protein; Tat protein;  
 KW vaccine; HIV infection; protein D.  
 XX  
 OS Synthetic.  
 OS Human immunodeficiency virus type 1.  
 XX  
 PN WO9916884-A1.  
 XX  
 PD 08-APR-1999.  
 XX  
 PF 17-SEP-1998; 98WO-EP06040.  
 XX  
 PR 26-SEP-1997; 97GB-0020585.  
 XX

(SMIK ) SMITHKLINE BEECHAM BIOLOGICALS.  
 PA Bruck C, Godart SAG, Marchand M;  
 PI  
 PI  
 XX  
 DR WPI; 1999-302282/25.  
 DR N-PSDB; AAX35688.  
 XX  
 XX HIV Tat or Nef protein linked to a fusion partner  
 XX  
 PS Disclosure; Fig 2; 66pp; English.  
 XX  
 CC The present sequence represents a fusion protein comprising LipD-HIV-1  
 CC Nef-His. The protein is exemplified by the fusion proteins of  
 CC the invention. The specification also describes fusion proteins  
 CC comprising HIV-1 Tat protein. The fusion protein can be used in a  
 CC vaccine to prevent HIV infection.  
 XX  
 SQ Sequence 324 AA;

Query Match 75.7%; Score 1706.5; DB 20; Length 324;  
 Best Local Similarity 78.8%; Pred. No. 1.6e-153;  
 Matches 324; Conservative 0; Mismatches 0; Indels 87; Gaps 1;  
 QY 1 CSHSSNMANTQMSDKIIIAHAGSGLPEHTLESKALFAQAQADLEQDLAMTKDGR 60  
 Db 1 cshssnmantqmsdkiiiahgagsglypehtleskalafagaqadyleqdlamtkdgrl 60  
 QY 61 VVIHDFLDGTVAKFPHRHRKDGRIYVIDFTLKEIQSLWENTENETGCGKWSKSVV 120  
 Db 61 vviindhldgtdvakkfphrhrkdgryyvidftlkelqslwemtentetmggkwsksvv 120  
 QY 121 GWPTVREMRRAEPAADGVGAASRDLEKHGAITSNTAATNAACAWLEAQEEVEGFPVT 180  
 Db 121 gwptvrermraepaadvggaasrdlekhgaitssntaatnaacawleaqeevevgfpvt 180  
 QY 181 PQVPLRPMYKAAVDLSHFLKEKGLGLIHSQRQDILDWYHTQGYFPDMQNYTPGP 240  
 Db 181 pqvplrptmykaavdlsfhflkekglglisqrqdlldwiyhtqgyfpdmqnytpgp 240

QY 241 GVRYPFTGWCYKLVPEPKVEANKGENTSLHPVSLHGMDDPEREVLWRFDLSRLAF 300  
 Db 241 gvrpyftfgwcyklvppepkveeankgentsllhpvslhgmddperewlwrfdslrlaf 300  
 QY 301 HHVARELHPEYFKNCTSEPVDRLEPWKHGSPQKPTACTNCTCYCKKCCFHCQVCFITKALG 360  
 Db 301 hhvarelhpeyfkntsepvdrlepwkhgspqktactnctcyckkccfhcqvcfitkalg 360  
 QY 361 ISYGRKKRRRPPQSGTHQVSLSKOPTSOSRGDPTGPKETSGHHHHH 411  
 Db 361 isygrkkrrrrppqsgthqvskskptdsrgdptgpketsghhhhh 411

RESULT 6  
 AAG63234  
 ID AAG63234 standard; Protein; 324 AA.  
 XX  
 AC AAG63234;  
 DT 01-OCT-2001 (first entry)  
 XX  
 DE Amino acid sequence of a His tagged LipD-Nef of HIV.  
 XX  
 KW HIV; nef gene; tat gene; vaccine; gp120 gene; HIV viral load.  
 XX  
 OS Synthetic.  
 OS Human immunodeficiency virus.  
 XX  
 FH Key Location/Qualifiers  
 FT Peptide 1..109  
 ET /note= "protD fusion partner"  
 XX  
 PN WO200154719-A2.  
 XX  
 PD 02-AUG-2001.  
 XX  
 PF 29-JAN-2001; 2001WO-EP00944.  
 XX  
 PR 31-JAN-2000; 2000GB-0002200.  
 PR 14-APR-2000; 2000GB-0009336.  
 PR 06-JUN-2000; 2000GB-0013806.  
 PR 28-JUN-2000; 2000WO-EP05998.  
 XX  
 PA (SMIK ) SMITHKLINE BEECHAM BIOLOGICALS.  
 XX  
 XX Voss G;  
 XX  
 XX WPI; 2001-476172/51.  
 DR N-PSDB; AAH42878.  
 XX  
 PT New use of a human immunodeficiency virus (HIV) Tat, or Nef, or Nef  
 PT linked to Tat (Nef-Tat) protein or polynucleotide and an HIV gp120  
 PT protein or polynucleotide for the manufacture of a vaccine -  
 PS Disclosure; Fig 1; 90pp; English.  
 XX  
 CC The present sequence represents a His-tagged Nef protein of HIV, with  
 CC a lipidation signal sequence (lipod) which is removed after processing.  
 CC The protein is expressed in Escherichia coli, and is used to  
 CC produce the vaccine of the invention. The specification describes  
 CC the use of HIV Tat, HIV Nef, or Nef-Tat; and HIV gp120 in the  
 CC manufacture of a vaccine. The vaccine is used for the prophylactic or  
 CC therapeutic immunization of humans against HIV. Tat, Nef or Nef-Tat act  
 CC in synergy with gp120 in the treatment and prevention of HIV. The  
 CC vaccine reduces the HIV viral load in HIV infected humans and results  
 CC in a maintenance of CD4+ levels over those levels found in the absence  
 CC of vaccination with HIV Tat, Nef or Nef-Tat and HIV gp120.  
 XX  
 SQ Sequence 324 AA;

Query Match 75.7%; Score 1706.5; DB 22; Length 324;







PR 31-JAN-2000; 2000GB-0002200.  
PR 14-APR-2000; 2000GB-0009336.  
PR 06-JUN-2000; 2000GB-0013806.  
PR 28-JUN-2000; 2000WO-EP05998.  
XX  
PA (SMIK ) SMITHKLINE BEECHAM BIOLOGICALS.

**Voss G;**

WPI: 2001-476172/51.

N-PSDB; AAH42877.

## New use of a human

protein or polynu

Disclosure: Fig 1: 90pp; English.

The present sequence represents a His-tagged Nef-Tat linked protein of HIV. The protein is expressed in the yeast *Pichia pastoris*, and is used to produce the vaccine of the invention. The specification describes the use of HIV Tat, HIV Nef, or Nef-Tat; and HIV gp120 in the manufacture of a vaccine. The vaccine is used for the prophylactic or therapeutic immunization of humans against HIV. Tat, Nef or Nef-Tat act in synergy with gp120 in the treatment and prevention of HIV. The vaccine reduces the HIV viral load in HIV infected humans and results in a maintenance of CD4+ levels over those levels found in the absence of vaccination with HIV Tat, Nef or Nef-Tat and HIV gp120.

Sequence 302 AA:

Query Match 74.9%; Score 1688; DB 22; Length 302;  
Best Local Similarity 100.0%; Pred. No. 8.2e-152;  
Matches 302: Conservative 0; Mismatches 0; Indels 0

Qy	110	MGSKWSSKSVGVHPTVRMRRAEPAAADGVGAASRDLEKHGAITSSNTAATNAACAWLEA	169
Db	1	mggkwsksvvgvptvrmrtraepaadgvgaasrdlekhgaitsntaataacswlea	60
Qy	170	QEEEEGVFPVTPQVPLRPMTYKAAVDLSHFLEKEKGLLEGIHSQRQDILDLDWIYHTQGY	229
Db	61	qeeeevgfpvtpqvpplrmtcykaavdlsfhlekeggleghlnsrqgdildldwihtqgy	120
Qy	230	FPQWQNTPGGVRYPDLTFCWCYKLYVPEDPKVEEANKGENTSLHPVSLHGMDDEREV	289
Db	121	fpdwqnytpgpgvryptftgwcylvpepdkveeankgentsllhpvslhgmdpber	180
Qy	290	LEWRFDRLAFHVARELHPYEPFKNCTSEVPDRLPELPWKHPGSQPTACTNCYCKKCCFH	349
Db	181	lewrfdrlafhvarelhpkyfknctsevpdrlepwhpgsqptactncycckccfh	240
Qy	350	COVCFITKALGISYGRKKRQRERRPQGSOTHOVSLSKQPTQSQRDPTGPKETSQHNNH	409
Db	241	cqvcfclkalgisyrkkrrqrrppqgsqthqvsksqptsgrgdpgpketsgnhh	300
Qy	410	HH 411	
Db	301	hh 302	

RESULT 11

AAY02357  
 ID AAY02357 standard; Protein; 302 AA.

XX  
AC AAY02357:

XX  
DE 00-777-1000, 1912-1913

XX

XX HIV nef gene; fuslon protein; HIV nef gene; Nef protein; Tat protein;  
KW vaccine; HIV infection; protein D.  
KW  
XX in representative HIV-1 mutant tat HIV protein.

XX	Synthetic.
OS	Human immunodeficiency virus type 1.
XX	
XX	WO9916884-A1.
XX	
XX	08-APR-1999.
PD	
XX	
XX	17-SEP-1998; 98WO-EP06040.
PF	
XX	
XX	26-SEP-1997; 97GB-0020585.
XX	
XX	(SMIK ) SMITHKLINE BEECHAM BIOLOGICALS.
PA	
XX	Bruck C, Goëart SAG, Marchand M;
PI	
XX	WPI: 1999-302282/25.
XX	N-PSDB: AAX35693
DR	

XX HIV Tat or Nef protein linked to a fusion partner  
PT  
XX  
PS Disclosure: Fig 2: 66pp; English.

xx The present sequence represents a representative HIV-1 mutant Tat-HIs  
cc protein. The protein is used in the creation of the fusion proteins of  
cc the invention, in conjunction with a fusion partner (e.g. protein D).  
cc The specification also describes fusion proteins comprising HIV-1 Nef  
cc protein. The fusion protein can be used in a vaccine to prevent HIV  
cc infection.

XX	Sequence	302 AA;
SQ		

```
Query Match - 74.3%; Score 1675; DB 20; Length 302;
Best Local Similarity 99.0%; Pred. No. 1.4e-150;
Matches 299; Conservative 2; Mismatches 1; Indels 0; Gaps 0;
```

Qy	110	MGGKWSKSSVVGNPTVPRVRMRRAEPAADGVGAASRDLEKHGATISNTAATNAACAWLEA	169
Db	1	mggkwsksvvgnptvrvmrrraepaadvggaasrdlekhgaitSENTAATNAACAWLEA	60
Qy	170	QEESVGFVPTPQVPLRPMTYKAAVDLSHFLKBEKGLLELIHSORRQDILDLIWYHTQGY	229
Db	61	qeesevfvptpqvplrpmtykaavdlsflkbekgleglihsqrrqdlldliwyhtqgy	120
Qy	230	FPDQNTPPGGVRYPLTFCWCYKLVPEPDKVPEEANKGENTSLLHPVSLHGMDPEREV	289
Db	121	fpdqntppggvrypltfcwcyklvpepdkvpeeankgentsllhpvs.lhgmdperev	180
Qy	290	LEWRFDSRLAFHHVARELHPEYFNKCTSEPVDRPLEPFWKHGPSQPTACTNCYCKKCCFH	349
Db	181	lewrfdsrlafhhvarelhpeyfnkctsepvdrplepwhhgsgqptactncycckckcfh	240
Qy	350	COVCFITKALGISYGRKKRORRPPGSGTOHVSLSKOPTSOSRGDPTGPKETSGHHNH	409
Db	241	cqvfcfkaalgisyrkrrqrrppgsgtqhvslskptsksgkgep9pketsgnhhh	300
Qy	410	HH 411	
Db	301	hh 302	

RESULT 12

AAG63239  
ID AAG63239 standard; Protein; 302 AA.

XX AAG63239:

XX  
XX

xx Amino acid sequence of a His-tagged mutant His protein of HIV.  
DE  
xx HIV: nef gene: rat gene: vaccine: gp120 gene: HIV viral load.  
KW

```

XX OS Synthetic.
OS Human immunodeficiency virus.
XX PN WO200154719-A2.
XX PD 02-AUG-2001.
XX PF 29-JAN-2001; 2001WO-EP00944.
XX PR 31-JAN-2000; 2000GB-0002200.
PR 14-APR-2000; 2000GB-0009336.
PR 06-JUN-2000; 2000GB-0013806.
PR 28-JUN-2000; 2000WO-EP05998.
XX PA (SMIK ) SMITHKLINE BEECHAM BIOLOGICALS.
XX PI Voss G;
XX DR WPI: 2001-476172/51.
DR N-PSDB; AAH42883.
XX PT New use of a human immunodeficiency virus (HIV) Tat, or Nef, or Nef
PT linked to Tat (Nef-Tat) protein or polynucleotide and an HIV gp120
PT protein or polynucleotide for the manufacture of a vaccine .
XX PS Disclosure; Fig 1; 90pp; English.
XX CC The present sequence represents a His-tagged mutant His protein of HIV.
CC The protein is expressed in the yeast Pichia pastoris, and is used to
CC produce the vaccine of the invention. The specification describes
CC the use of HIV Tat, HIV Nef, or Nef-Tat; and HIV gp120 in the
CC manufacture of a vaccine. The vaccine is used for the prophylactic or
CC therapeutic immunization of humans against HIV. Tat, Nef or Nef-Tat act
CC in synergy with gp120 in the treatment and prevention of HIV. The
CC vaccine reduces the HIV viral load in HIV infected humans and results
CC in a maintenance of CD4+ levels over those levels found in the absence
CC of vaccination with HIV Tat, Nef or Nef-Tat and HIV gp120.
XX SQ Sequence 302 AA:

Query Match 74.3%; Score 1675; DB 22; Length 302;
Best Local Similarity 99.0%; Pred. No. 1.4e-150;
Matches 299; Conservative 2; Mismatches 1; Indels 0; Gaps 0;

QY 110 MGGKWSKSSVVGWPTVRMRRAEPAADGVGAASRDLEKHGAISSNTAATNAACAWLEA 169
DB 1 mggkwsksvvgwptvrermraepaadvgaaardlekghaltsntaatnaacawlea 60

QY 170 QEEEEVGFPVTPQVPLRPMTYKAAVDLSHFLKEKGGLEGILHSQRRQDILDLIWYHTQGY 229
DB 61 qeeeevgfpvtpqvpplrpmtykaavdlshflkekgglegilhsqrrqdlldliwyhtqgy 120

QY 230 FPDQWNTPGGVRYPLTFGWCYKLVPEPDKVEEANKGENTSLLHPVSLHGMDPEREV 289
DB 121 fpdwqntpggvrypltfgcwcyklyvpdpdkveeankgentsllhpvs lhgmdperev 180

QY 290 LEWRFDSRLAFHVAHELHPEYFNKNTSEPVDPRLPEPKHPGQPKTACTNCYCKCCFFH 349
DB 181 lewrfdsrlafhvharelhpeyfnkntsepvdp rlepkhp gspktactncyc kccffh 240

QY 350 CQVCFITKALGISYGRKKRRRPPQGSQTHQVSLSKQPTSOSRGDPTGPKETSGHHHH 409
DB 241 cqvcfitaalgisygrkkrrppqgsqthqvs lskqptsosrgdptgpketsghhhh 300

QY 410 HH 411
DB 301 hh 302

RESULT 13
AAV02349

```

```

ID AAV02349 standard; Protein; 215 AA.
XX AC AAV02349;
XX DT 09-JUL-1999 (first entry)
XX DE A representative HIV-1 Nef-His protein.
XX KW HIV nef gene; fusion protein; HIV nef gene; Nef protein; Tat protein;
XX OS Synthetic.
XX OS Human immunodeficiency virus type 1.
XX PN WO9916884-A1.
XX PD 08-APR-1999.
XX PF 17-SEP-1998; 98WO-EP06040.
XX PR 26-SEP-1997; 97GB-0020585.
XX PA (SMIK ) SMITHKLINE BEECHAM BIOLOGICALS.
XX PI Bruck C, Godart SAG, Marchand M;
XX DR WPI: 1999-302282/25.
DR N-PSDB; AAX35885.
XX PT HIV Tat or Nef protein linked to a fusion partner
XX PS Disclosure; Fig 2; 66pp; English.
XX CC The present sequence represents a representative HIV-1 Nef-His protein.
CC The protein is used in the creation of the fusion proteins of
CC the invention, in conjunction with a fusion partner (e.g. protein D).
CC The specification also describes fusion proteins comprising HIV-1 Tat
CC protein. The fusion protein can be used in a vaccine to prevent HIV
CC infection.
XX SQ Sequence 215 AA:

Query Match 50.5%; Score 1139.5; DB 20; Length 215;
Best Local Similarity 71.2%; Pred. No. 6.4e-100;
Matches 215; Conservative 0; Mismatches 0; Indels 87; Gaps 1;

QY 110 MGGKWSKSSVVGWPTVRMRRAEPAADGVGAASRDLEKHGAISSNTAATNAACAWLEA 169
DB 1 mggkwsksvvgwptvrermraepaadvgaaardlekghaltsntaatnaacawlea 60

QY 170 QEEEEVGFPVTPQVPLRPMTYKAAVDLSHFLKEKGGLEGILHSQRRQDILDLIWYHTQGY 229
DB 61 qeeeevgfpvtpqvpplrpmtykaavdlshflkekgglegilhsqrrqdlldliwyhtqgy 120

QY 230 FPDQWNTPGGVRYPLTFGWCYKLVPEPDKVEEANKGENTSLLHPVSLHGMDPEREV 289
DB 121 fpdwqntpggvrypltfgcwcyklyvpdpdkveeankgentsllhpvs lhgmdperev 180

QY 290 LEWRFDSRLAFHVAHELHPEYFNKNTSEPVDPRLPEPKHPGQPKTACTNCYCKCCFFH 349
DB 181 lewrfdsrlafhvharelhpeyfnkntsepvdp rlepkhp gspktactncyc kccffh 206

QY 350 CQVCFITKALGISYGRKKRRRPPQGSQTHQVSLSKQPTSOSRGDPTGPKETSGHHHH 409
DB 207 -----tsghhhh 213

QY 410 HH 411
DB 214 hh 215

RESULT 14

```

```

Db      207  -----|||
Qy      410  HH 411
        ||
Db      214  hh 215

RESULT 15
AAY50795
ID      AAY50795 standard; protein; 206 AA.
XX
XX      AAY50795;
XX
XX      17-FEB-2000 (first entry)
XX
XX      Human NEF protein/calmodulin binding inhibitor.
XX
XX      NEF protein; calmodulin; negative factor protein; binding inhibitor;
XX      diagnosis; detection; infection; treatment; HIV.
XX
XX      Homo sapiens.
XX
XX      WO9957136-A2.
XX
XX      11-NOV-1999.
XX
XX      06-MAY-1999; 99WO-EP03105.
XX
XX      06-MAY-1998; 98DE-1020224.
XX
XX      (SCHO/) SCHOTT M.
XX      (SCHO/) SCHORR J.
XX      (ANTZ/) ANTZ C.
XX
XX
XX      Schott M, Schorr J, Antz C;
XX      WPI; 2000-038789/03.
XX
XX      Binding agents used for treatment, prevention and diagnosis of human
XX      immune deficiency virus infection
XX
XX      Claim 3; Fig 1; 29pp; German.
XX
XX      This invention describes novel binding partners (A) (i) for negative
XX      factor protein (Nef) that competitively inhibit binding of calmodulin (i)
XX      to Nef, or (ii) for (i) that competitively inhibit binding of Nef to (i).
XX      (A) are used for prevention, diagnosis (by specific detection of intra-
XX      and/or extra-cellular Nef, including staging of infection), and/or
XX      treatment of infections with human immune deficiency virus (HIV).
XX
XX      Sequence 206 AA;
XX
XX      Query Match 49.7%; Score 1120; DB 21; Length 206;
XX      Best Local Similarity 99.0%; Pred. No. 4.3e-98;
XX      Matches 204; Conservative 0; Mismatches 2; Indels 0; Gaps 0;
Qy      110  MGGKWSKSSVVGWPTVRERMRRAEPAADGVGAASRDLEKKGAITSSNTAATNAACAWLEA 169
Db      1    |mggkwsksvvvgwpavrmrrraepaagvgaaasrdlekkgaitssntaannaacawlea 60
Qy      170  QEEEEVGPVTPQVPLRPMTYKAAVDLSHFLKEKGGLEGSLHSQRRQDILDWIYHTQGY 229
Db      1    |qeeeevgfpvtpqvpplrmpykaaavdlshflkekgglegslhsqrqrqldlwiyhtggy 120
Qy      230  FPDQWNYTPGCVRYPLTFGWCYKLVPEVPDKVEEANKGENTSLILHPVSLHGMDDPEREV 289
Db      1    |fpdqwnytpgcvrypltfgwcylvpevpdkveeankgentsllhpvslhgmddperev 180
Qy      290  LEWRFDSRLAFHHVARELHPEYFKNC 315
Db      181  lewrfdsrlafhhvarelhpvfyknc 206

```

## RESULT 16

AAR38893  
ID AAR38893 standard; Protein; 206 AA.

XX AC AAR38893;  
XX DT 10-NOV-1993 (first entry)  
XX DE Nef protein of HIV-1.  
XX KW AIDS: antibody; p25; gp110; gp41; assay; detection;  
XX KW immunity; vaccine.  
XX OS Human immunodeficiency virus-1.

XX PN US221610-A.  
XX PD 22-JUN-1993.  
XX PF 26-MAY-1988; 88US-0199143.  
XX PR 26-MAY-1988; 88US-0199143.  
XX PR 04-SEP-1991; 91US-0754300.

XX PA (INEM) INST NAT SANTE & RECH MEDICALE.  
XX PA (INSP) INST PASTEUR.  
XX PI Bahraoui EM, Chamaret S, Ferris S, Granier C, Montagnier L;  
XX PI Rietschoten JV, Rochat H, Sabatier JM;  
XX DR WPI; 1993-213434/26.  
XX PT Diagnosis of HIV infection - by detecting HIV antibodies using  
XX PT antigenic polypeptide derived from nef protein of HIV-1  
XX PS Disclosure; Fig 2; 15pp; English.

XX CC The nef protein comprises peptides which are expressed in vivo in HIV  
XX CC infected patients before detectable amts. of p25, gp110 and gp41 are  
XX CC expressed. Thus, they can be used in assays for early detection of HIV.  
XX CC They can also be used to raise antibodies for use in detection,  
XX CC to induce cellular immunity or to raise neutralising antibodies  
XX CC that either inactivate the AIDS virus or reduce the viability of  
XX CC the virus in vivo or destroy infected cells.  
XX CC The peptides may be used in viral vaccines.

XX SQ Sequence 206 AA;

Query Match 49.5%; Score 1116; DB 14; Length 206;  
Best Local Similarity 99.0%; Pred. No. 1e-97;  
Matches 204; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 110 MGGKSKSSVVGWPTVRMRRAEPAADGVGAASRDLEKHGALTSSNTAATNAACAWLEA 169  
DB 1 mggkssksvvgwptvrmmraepaadvgaasrdlekhgaltssntaataacawlea 60  
QY 170 QEEEEVGPVTPQVPLRPMTYKAAVDLSHFLKEKGLEGLIHSQRQDILDLMVHTQGY 229  
DB 61 qeeeevfpvtqpvlrmpmtkaavdlshflkekgleglshqrqrdildlvihtqgy 120  
QY 230 FPDWQNYTPGPGVRYPLTFCWCYKLVPEPDKVEANKGENTSLHPVSLHGMDDPEREV 289  
DB 121 fpdwnytpgpgvrypltfwcycylvpdpdkveeankgentsllhpvslhgmddperev 180  
QY 290 LEWRFDLSRLAFHHVARELHPYFKNC 315  
DB 181 lewrfdslafhhvarelhpeyfknc 206

## RESULT 17

## AAB10054

ID AAB10054 standard; Protein; 206 AA.

XX AC AAB10054;  
XX DT 02-NOV-2000 (first entry)  
XX DE HIV-1 nef protein.  
XX KW Glycoprotein; gag; pol; GP-1; GP-2; anti-HIV; cytostatic;  
XX KW gene therapy; treatment; infectious disease; HIV; AIDS; neoplasm;  
XX KW carcinoma; melanoma; nef protein.  
XX OS Human immunodeficiency virus type 1.

XX PN EP1006196-A2.  
XX PD 07-JUN-2000.  
XX PF 25-NOV-1999; 99EP-0250415.  
XX PR 26-NOV-1998; 98DE-1056463.  
XX PA (PETT-) PETTE INST HEINRICH.  
XX PI Von Laer MD;  
XX DR WPI; 2000-378268/33.  
XX DR N-PSDB; AAA40298, AAB10053, AAB10054.

XX PT New retroviral packing cell useful as pharmaceutical carrier in gene  
XX PT therapy for treatment of HIV and neoplasms, comprises retroviral genes  
XX PT and glycoproteins -  
XX PS Disclosure: Page 48; 69pp; German.  
XX CC This invention describes a novel retroviral packing cell (I), comprising  
XX CC the retroviral genes gag, pol and glycoproteins GP-1 and GP-2 of the LCMV  
XX CC coding gene gp, or a part of these. The products of the invention have  
XX CC anti-HIV and cytostatic activity and can be used for gene therapy. (I) is  
XX CC useful for in vitro infection of cells, especially hematopoietic stem  
XX CC cells, for expression of transgenes in cells and as a pharmaceutical  
XX CC carrier for gene therapy. (I) is therefore useful in the treatment of  
XX CC infectious diseases (HIV or AIDS), neoplasms (carcinomas), melanomas and  
XX CC other diseases. This sequence represents the Human immunodeficiency virus  
XX CC (HIV-1) nef protein described in the method of the invention.

XX SQ Sequence 206 AA;

Query Match 49.4%; Score 1115; DB 21; Length 206;  
Best Local Similarity 98.1%; Pred. No. 1.3e-97;  
Matches 202; Conservative 1; Mismatches 3; Indels 0; Gaps 0;

QY 110 MGGKSKSSVVGWPTVRMRRAEPAADGVGAASRDLEKHGALTSSNTAATNAACAWLEA 169  
DB 1 mggkssksvvgwptvrmmraepaadvgaasrdlekhgaltssntaanaacawlea 60  
QY 170 QEEEEVGPVTPQVPLRPMTYKAAVDLSHFLKEKGLEGLIHSQRQDILDLMVHTQGY 229  
DB 61 qeeeevfpvtqpvlrmpmtkaavdlshflkekgleglshqrqrdildlvihtqgy 120  
QY 230 FPDWQNYTPGPGVRYPLTFCWCYKLVPEPDKVEANKGENTSLHPVSLHGMDDPEREV 289  
DB 121 fpdwnytpgpgvrypltfwcycylvpdpdkveeankgentsllhpvslhgmddperev 180  
QY 290 LEWRFDLSRLAFHHVARELHPYFKNC 315  
DB 181 lewrfdslafhhvarelhpeyfknc 206

RESULT 18  
AAW89326

ID AAW89326 standard; Protein; 206 AA.  
XX  
AC AAW89326;  
XX  
DT 01-JUN-1999 (first entry)  
XX  
DE HIV-1 nef protein sequence.  
XX  
KW Antigenic composition; primate; lentivirus; nef gene; vaccine;  
KW infection; AIDS; HIV-1; nef protein.  
XX  
OS Human immunodeficiency virus type 1.  
XX  
PN US5851813-A.  
XX  
PD 22-DEC-1998.  
XX  
PF 27-JAN-1994; 94US-0188583.  
XX  
PR 27-JAN-1994; 94US-0188583.  
PR 12-JUL-1990; 90US-0551945.  
PR 09-JUL-1991; 91US-0727494.  
XX  
PA (HARD ) HARVARD COLLEGE.  
XX  
PI Destosiers RC;  
XX  
PS WPI; 1999-080408/07.  
XX  
DR N-PSDB; AAW81866.  
XX  
PT Lentivirus antigenic compositions - containing lentivirus with nef  
XX gene deletion  
XX  
PS Disclosure; Fig 2A-R; 93pp; English.  
XX  
CC The invention relates to an antigenic composition comprising an isolated  
CC primate lentivirus whose genome contains an engineered non-revertible  
CC null mutation in the nef gene, or an infectious DNA clone in a carrier.  
CC The antigenic composition is used in vaccines against infection by the  
CC lentivirus, e.g. AIDS.  
XX  
SQ Sequence 206 AA;  
  
Query Match 49.4%; Score 1114; DB 20; Length 206;  
Best Local Similarity 97.6%; Pred. No. 1.6e-97;  
Matches 201; Conservative 3; Mismatches 2; Indels 0; Gaps 0;  
  
QY 110 MGKWSKSSVGVGPTVRMRRAEPAADGVGAASRDLKKGALTSSNTAATNAACAWLEA 169  
Db 1 mggkwsksvlgvpavremrrraepaadvgaasrdlekghaltsntaannaacawlea 60  
  
QY 170 QEEEEVGFVPTQVPLRPMYKAAVDLSHFLKEGGLIHSQRQDILDLMYHTQGY 229  
Db 61 qeeekvgfvtqvpvlpmtykaavdlshflkekgleglilhsqrqdlldlwyhtqgy 120  
  
QY 230 FPDQNYTFPGVRYPLTFGWCKYKLVPEPDKVEEANKGENTSLLHPVSLHGMDPEREV 289  
Db 121 fpdwqnytpgpgirypvltfgwcyklypvpdkveeankgentsllhpvsilhgmddperv 180  
  
QY 290 LEWRFDRLAFHHVARELHPEYFKNC 315  
Db 181 lewrfdslafhhvarelhpeyfknc 206  
  
RESULT 19  
AAP61515  
ID AAP61515 standard; Protein; 206 AA.  
XX  
AC AAP61515;  
XX  
DT 08-JUN-1991 (first entry)  
XX

DE Sequence of E' protein.  
XX  
XX HIV; LAV; AIDS; diagnosis; vaccine.  
XX  
OS HTLV-IIIB/H9 cells (ATCC CRL 8543).  
XX  
PN EP187041-A.  
XX  
PD 09-JUL-1986.  
XX  
PF 23-DEC-1985; 85EP-0309454.  
XX  
PR 24-DEC-1984; 84US-0685272.  
PR 04-DEC-1985; 85US-0805069.  
XX  
PA (GETH ) GENENTECH INC.  
XX  
PI Capon DJ, Lasky LA;  
XX  
DR WPI; 1986-177602/28.  
DR N-PSDB; AAN60288.  
XX  
PT Acquired immune deficiency syndrome polypeptide(s) - obt'd. by  
PT molecular cloning etc. and used for diagnosis and in vaccines  
PT against virus disease  
XX  
PS Example; fig 2; 125pp; English.  
XX  
CC A comparison of AAN60287 with the cDNA of the HTLV-III genome  
CC revealed one particular clone, designated p7.11 which contained a  
CC DNA sequence encoding this peptide (AAP60308) sequence. This approx.  
CC 2.2 kilobase covers the precursor gag region and encodes, 5' to 3',  
CC p-12, p-15, p-24 a second p-15 protein, and approx. 300 extra base  
CC pairs 3' to the gag region (see AAN60288).  
XX  
SQ Sequence 206 AA;  
  
Query Match 49.3%; Score 1111; DB 7; Length 206;  
Best Local Similarity 97.1%; Pred. No. 3.1e-97;  
Matches 200; Conservative 4; Mismatches 2; Indels 0; Gaps 0;  
  
QY 110 MGKWSKSSVGVGPTVRMRRAEPAADGVGAASRDLKKGALTSSNTAATNAACAWLEA 169  
Db 1 mggkwsksvlgvpavremrrraepaadvgaasrdlekghaltsntaannaacawlea 60  
  
QY 170 QEEEEVGFVPTQVPLRPMYKAAVDLSHFLKEGGLIHSQRQDILDLMYHTQGY 229  
Db 61 qeeekvgfvtqvpvlpmtykaavdlshflkekgleglilhsqrqdlldlwyhtqgy 120  
  
QY 230 FPDQNYTFPGVRYPLTFGWCKYKLVPEPDKVEEANKGENTSLLHPVSLHGMDPEREV 289  
Db 121 fpdwqnytpgpgirypvltfgwcyklypvpdkveeankgentsllhpvtlhgmddperv 180  
  
QY 290 LEWRFDRLAFHHVARELHPEYFKNC 315  
Db 181 lewrfdslafhhvarelhpeyfknc 206  
  
RESULT 20  
AAP60423  
ID AAP60423 standard; Protein; 216 AA.  
XX  
AC AAP60423;  
XX  
DT 20-AUG-1991 (first entry)  
XX  
DE Sequence of LAV virus ORF F protein.  
XX  
KW AIDS vaccine; diagnosis; immunoassay; HIV; HTLV-III.  
XX  
OS Lymphadenopathy virus.  
XX

PN WO8602383-A.  
 XX  
 PD 24-APR-1986.  
 XX  
 XX 18-OCT-1985; 85WO-EP00548.  
 XX  
 XX 21-JAN-1985; 85GB-0001473.  
 PR 18-OCT-1984; 84FR-0016013.  
 PR 16-NOV-1984; 84GB-0029099.  
 XX  
 XX (CNRS ) CNRS CENT NAT RECH SCI.  
 PA (INSP ) INST PASTEUR.  
 XX  
 XX Montagnier L, Krust B, Chamaret S, Clavel F, Chermann J-C;  
 PI Barre-Sinoussi F, Allison M, Sonigo P, Stewart C, Danos O;  
 PI Wain-Hobson S;  
 XX  
 XX WPI: 1986-119166/18.  
 DR N-PSDB; AAN60365.  
 XX  
 XX Purified glyco:protein and peptide(s) - are recognised by sera contg.  
 PT antibodies against lymphadenopathy virus and useful in detecting  
 PT AIDS antibodies or in vaccines  
 XX  
 XX Disclosure; Fig 4; 75pp; English.  
 XX  
 CC The inventors claim a polypeptide which is recognised by sera of  
 CC human origin contg. antibodies against the virus of  
 CC lymphadenopathies (LAV) or acquired immune deficiency syndrome  
 CC (AIDS). Also claimed are various peptides corresp. to the AA  
 CC sequences deducible from proteins encoded by LAV DNA, defined by  
 CC specific residues (e.g. 12-32, 37-46, 49-79, 88-153) in accordance  
 CC with a formula given in the specification.  
 XX  
 XX Sequence 216 AA;  
 SQ

Query Match 49.1%; Score 1107; DB 7; Length 216;  
 Best Local Similarity 98.1%; Pred. No. 7.9e-97;  
 Matches 202; Conservative 1; Mismatches 3; Indels 0; Gaps 0;

QY 110 MGGKWSKSSVGVGPTVREMRRAEPAADGVGAASRDLEKHGATSSNTAATNAACAWLEA 169  
 DB 11 mggkwsksvvgvptvremtraepaahvgvaaspdekhgaltssntaataacawlea 70  
 QY 170 QEEEEVGFVPTQVPLRPMTYKAADVLSHFLKEKGLGGLIHSQRQDILDWYHTQGY 229  
 DB 71 qeeevgfvptqvpplrpmtykaadvlsfhfkekgglgglisqrqdlldwlyhtqgy 130  
 QY 230 FPDQNYTPGPGVRYPLTFGWCYKLYPVPDPKVEEANKGENTSLHPVSLHGMDPEREV 289  
 DB 131 fpdqnytpgpgvrypltfwgcyklypvpdpkveeankgentsllhpvslhgmdperv 190  
 QY 290 LEWRFDSRLAFHVAHELHPYFKNC 315  
 DB 191 lewrfdsrlafhvharelhpqyfknc 216

RESULT 21  
 AAW90179  
 ID AAW90179 standard; Protein; 206 AA.  
 XX  
 AC AAW90179;  
 XX  
 DT 14-JUN-1999 (first entry)  
 XX  
 DE HTLV-III E' protein.  
 XX  
 KW AIDS; acquired immunodeficiency syndrome; viral infection; env protein;  
 KW fusion polypeptide; E' protein; gag protein; pol protein; p' protein;  
 KW gp signal peptide; detection; vaccination; etiological agent;  
 KW infection inhibitor; AIDS associated retrovirus; p24; gp41.  
 XX

OS Human lymphotropic virus type III.  
 XX  
 FH Key Location/Qualifiers  
 FT Misc-difference 11 /label= Val, Ile  
 FT Misc-difference 51 /label= Thr, Asn  
 FT Misc-difference 54 /label= Asp, Ala  
 FT Misc-difference 65 /label= Glu, Lys  
 FT Misc-difference 151 /label= Glu, Asp  
 FT Misc-difference 153 /label= Leu, Val  
 XX  
 PN US5853978-A.  
 XX  
 PD 29-DEC-1998.  
 XX  
 PF 29-JUL-1994; 94US-0282857.  
 XX  
 PR 08-MAY-1986; 86US-0861016.  
 PR 24-DEC-1984; 84US-0685272.  
 PR 04-DEC-1985; 85US-0805069.  
 PR 02-AUG-1988; 88US-0227568.  
 PR 19-NOV-1992; 92US-0979391.  
 PR 29-SEP-1993; 93US-0129009.  
 PR 29-JUL-1994; 94US-0282857.  
 XX  
 PA (GETH ) GENENTECH INC.  
 XX  
 XX Bertran PW, Capon DJ, Lasky LA;  
 PI  
 DR WPI: 1999-094894/08.  
 DR N-PSDB; AAV74271.  
 XX  
 XX New AIDS-associated fusion poly:peptide(s) - used for the detection  
 PT of AIDS or to inhibit infection by AIDS associated retrovirus or  
 PT dissemination of such retrovirus in infected individuals  
 XX  
 XX Example 1; Fig 2A-J; 47pp; English.  
 PS  
 CC This invention describes a novel fusion polypeptide which comprises  
 CC (a) a first polypeptide sequence of an AIDS associated E', env, or gag  
 CC polypeptide, that specifically binds complementary antibody and (b) a  
 CC second polypeptide sequence which is not an AIDS associated virus  
 CC polypeptide. Also described are (1) a fusion polypeptide having a  
 CC herpes simplex virus (HSV) gp signal peptide sequence fused in reading  
 CC frame with a polypeptide sequence other than HSV gp, (2) a nucleic acid  
 CC encoding a fusion polypeptide as in (1); (3) an expression vector  
 CC comprising a nucleic acid as in (2); (4) a host comprising a vector as  
 CC in (3). The AIDS-associated fusion polypeptides can be used in the  
 CC detection of and vaccination against viral etiological agents of AIDS.  
 CC They can also be administered as a pharmaceutical agent to inhibit  
 CC infection by AIDS associated retrovirus or dissemination of such  
 CC retrovirus in infected individuals.  
 XX  
 XX Sequence 206 AA;  
 SQ

Query Match 48.4%; Score 1092; DB 20; Length 206;  
 Best Local Similarity 96.1%; Pred. No. 2e-95;  
 Matches 198; Conservative 1; Mismatches 7; Indels 0; Gaps 0;

QY 110 MGGKWSKSSVGVGPTVREMRRAEPAADGVGAASRDLEKHGATSSNTAATNAACAWLEA 169  
 DB 1 mggkwsksvvgvptvremtraepaahvgvaaspdekhgaltssntaataacawlea 60  
 QY 170 QEEEEVGFVPTQVPLRPMTYKAADVLSHFLKEKGLGGLIHSQRQDILDWYHTQGY 229  
 DB 61 qeeevgfvptqvpplrpmtykaadvlsfhfkekgglgglisqrqdlldwlyhtqgy 120

QY 230 FPDQNYTPGVRVPLTFCWCYKLVPEVPDKVVEANKGENTSLHPVSLRGMDPEREV 289  
|||||  
Db 121 fpdqnytpgvriryltfgwcyklyvpvexpxeeankgentllhpsvhlhgmddperev 180  
  
QY 290 LEWRFSRLAFHHVARELHPEYFKNC 315  
|||||  
Db 181 lewrfsrlafhhvarelhpeyfknc 206  
  
RESULT 22  
AAE04960  
ID AAE04960 standard; Protein; 216 AA.  
AC AAE04960;  
XX  
DT 10-SEP-2001 (first entry)  
XX  
DE HIV-1 Jrf1 Nef protein.  
XX  
KW DNA vaccine; codon optimisation; Nef protein; HIV infection; anti-HIV;  
KW cell mediated immune response; cytotoxic T lymphocyte; CTL;  
KW human immunodeficiency virus-1; HIV-1; mutant; mutain.  
XX  
OS Human immunodeficiency virus type 1.  
XX Synthetic.  
XX  
FH Key Location/Qualifiers  
FT Modified-site 2 /note= "Myristylation site"  
XX  
XX WO200143693-A2.  
XX 21-JUN-2001.  
XX  
XX 15-DEC-2000; 2000WO-US34162.  
XX  
XX 17-DEC-1999; 99US-0172442.  
XX (MERI ) MERCK & CO INC.  
XX  
XX Shiver JW, Liang X, Fu T;  
XX  
XX WPI; 2001-417878/44.  
XX N-PSDB; AAD09603.  
XX  
XX Novel DNA vaccine useful for inducing cell mediated immune response  
PT against virulent strains of HIV, comprises DNA expression vector, and  
PT DNA containing codon optimized open reading frame encoding Nef protein  
PT  
XX  
XX Claim 5; Page 13; 84pp; English.  
XX  
XX The present invention relates to a DNA vaccine comprising a DNA  
CC expression vector, and a DNA molecule containing a codon optimised open  
CC reading frame encoding a Nef protein or its immunogenic derivative. The  
CC Nef protein or its derivative is expressed and generates an immune  
CC response which provides a substantial level of protection against HIV  
CC infection, upon administration of to a host. The DNA vaccine is useful  
CC for inducing a cell mediated immune (cytotoxic T lymphocyte, CTL)  
CC response against infection or disease caused by virulent strains of HIV,  
CC by administering the DNA vaccine into the tissue of the vertebrate host,  
CC preferably human. It is also useful for lowering transmission rate to  
CC previously uninfected individuals and/or for reducing levels of viral  
CC loads within an infected individual, so as to prolong the asymptomatic  
CC phase of HIV infection. The present sequence is human immunodeficiency  
CC virus-1 (HIV-1) jrf1 Nef protein. The different codon optimised nef  
CC constructs of the invention are used in the exemplification of the  
CC invention.  
XX  
XX Sequence 216 AA;  
SQ

Query Match 44.8%; Score 1011; DB 22; Length 216;

Best Local Similarity 83.8%; Pred. No. le-87;  
Matches 181; Conservative 15; Mismatches 10; Indels 10; Gaps 1;  
  
QY 110 MGKSKSVGWGPTVRMRRAEPAAD-----GVGAASRDLEKHGAITSNTAA 159  
|||||  
Db 1 mggkswksrvgwstvrermraepaadvrrtpeaavgvavsrdekhgaitssntaa 60  
  
QY 160 TNAACAWLEAQEEVEGFPVTPQVPLRPMTYKAAVDLSHFLKEKGLEGLIHSORQDIL 219  
|||||  
Db 61 tnaacawleaqeadeevgfpvrpqvplrpmtkygavdlshflkekggleglihsqkrqdl 120  
  
QY 220 DLMIYHTQGYFPDQNYTPGVRVPLTFCWCYKLVPEVPDKVVEANKGENTSLHPVSL 279  
|||||  
Db 121 dlwyhtqgyfpdwqnytpgvgirflptfgwcfklypvepekeveeanegenncllhpmqs 180  
  
QY 280 HCMDDPEREVLWRFSRLAFHHVARELHPEYFKNC 315  
|||||  
Db 181 hgiedpekevelewrfsrlafhhvarelhpeyfknc 216  
  
RESULT 23  
AAW53113  
ID AAW53113 standard; Protein; 210 AA.  
XX  
AC AAW53113;  
XX  
DT 25-JUN-1998 (first entry)  
XX  
DE Protein 6 contained in a complete ARV-2 nucleotide sequence.  
XX  
KW ARV-2; enhanced promoter; gene expression; cytomegalovirus;  
KW HIV; AIDS.  
XX  
OS Human Immunodeficiency virus type 1.  
XX  
XX US5688688-A.  
XX 18-NOV-1997.  
XX  
XX 10-AUG-1994; 94US-0288336.  
XX  
XX 24-DEC-1987; 87US-0138894.  
XX 31-OCT-1984; 84US-0667501.  
XX 30-JAN-1985; 85US-0696534.  
XX 06-SEP-1985; 85US-0773447.  
XX 17-AUG-1992; 92US-0931191.  
XX 28-JUN-1993; 93US-0083391.  
XX 17-AUG-1993; 93US-0107377.  
XX 10-AUG-1994; 94US-0288336.  
XX (CHIR ) CHIRON CORP.  
XX  
XX Chapman BS, Dina D, Haigwood NL, Luciw PA, Rosenberg S;  
XX Thayer RM;  
XX WPI; 1998-007982/01.  
XX  
XX Enhanced promoter for gene expression - comprising cytomegalovirus  
PT immediate early promoter plus intron  
XX  
XX Example 1; Fig 4C-P; 99pp; English.  
XX  
XX This sequence represents a protein of unspecified function contained in a  
CC complete nucleotide sequence of ARV-2 derived from partial sequences of  
CC several ARV clones. The invention provides a method for construction of a  
CC vector for expression of a polypeptide in a mammalian cell, comprising a  
CC polypeptide coding sequence operably linked downstream of an enhanced  
CC promoter. The enhanced promoter comprises the human cytomegalovirus  
CC immediate early region (HCMV IE1) promoter and the first intron proximate  
CC to the 3' end of the HCMV IE1 promoter. The polypeptide can be any of the  
CC HIV recombinant polypeptides and especially HIV gp120. Expression of HIV  
CC gp120 by COS 7 cells transfected with pCMW6a containing the gp120 coding  
CC region, where pCMW6a is a vector containing the above enhanced promoter,  
CC



CC is increased by a factor of 50-100 compared with the use of a vector  
CC containing the SV40 early promoter.

XX Sequence 210 AA;  
SQ Query Match 44.2%; Score 997; DB 19; Length 210;  
Best Local Similarity 85.2%; Pred. No. 2.1e-86;  
Matches 179; Conservative 16; Mismatches 11; Indels 4; Gaps 1;  
QY 110 MGGKWSKSSVVGWPTVYRMR-----RAEPAADGVGAASRDLEKHGAITSNTAATNAACA 165  
DB 1 mggkwsksrmggwsalrerrmrtaepaagdgavsdlekghgaitssntaataadca 60  
QY 166 WLEAQEEVEGFPVTPQVPLRPMYKAAVDLSHFLKEKGLEGLIHSQRQDILDLYIH 225  
DB 61 wleaqeeveegfpvtpqvlrmpmtykaaidishflkekgleglilwsqrqelldlyih 120  
QY 226 TQGYFPDMQNYTPGCVRYPLTFGWCYKLVPEPDKVEEANKGENTSLHLPVSLHGMDDP 285  
DB 121 tqgyfpdmqnytpgcvryplrtfgwciklvpepekeveeanegeenllhpmshgmeda 180  
QY 286 EREVLEWRFDSRLAFHHVARELHPEYFKNC 315  
DB 181 ekevlwrfdsklafhhmarelhpeyykdc 210

## RESULT 24

AAAY77299  
ID AAY77299 standard; Protein; 210 AA.  
XX AC AAY77299;  
XX DT 22-MAY-2000 (first entry)  
XX DE HIV-1 (ATCC CRL 8597) ORF #3-encoded protein.  
XX KW HIV-1; immunoassay; antigen; pol fragment; p31; recombinant; antibody;  
XX KW detection.

Human immunodeficiency virus type 1 'ATCC CRL 8597'.

US6013432-A.  
PD 11-JAN-2000.  
XX PF 17-MAY-1995; 95US-0443434.  
XX PR 08-JUL-1993; 93US-0089407.  
XX PR 24-DEC-1987; 87US-0138894.  
XX PR 17-AUG-1992; 92US-0931154.  
XX PR 31-OCT-1984; 84US-0667501.  
XX PR 30-JAN-1985; 85US-0696534.  
XX PR 06-SEP-1985; 85US-0773447.

(CHIR ) CHIRON CORP.

Luciw PA, Dina D;

WPI; 2000-170256/15.

DR N-PSDB; AA290201.

XX Immunoassay for antibodies against human immune deficiency virus, for  
PT diagnosing infection, uses an immunogenic fragment of the pol protein  
PT as antigen -

XX Example 1; Fig 40-P; 99pp; English.

XX The invention relates to the improvement of HIV-1 immunoassays by the  
CC use of an HIV-1 antigen comprising an immunogenic fragment of  
CC recombinant or synthetic HIV-1 pol, which is encoded by an approximately  
CC 9.7 kb sequence between a BstXI restriction site at position 3006 and an  
CC NdeI site at position 5131 of the genome (the proviral DNA sequence is

CC given in AA290201). The immunogenic pol fragment is not immunologically  
CC cross-reactive with human T cell lymphotropic viruses 1 or II. The  
CC invention also encompasses the use of p31 as an antigen. The recombinant  
CC antigens may be produced in Escherichia coli, Saccharomyces cerevisiae or  
CC in mammalian cells. Immunoassays using the recombinant HIV proteins may  
CC be used to diagnose and stage HIV-1 infections. Sequences  
CC AA77294-777299 represent proteins encoded by the genome of HIV-1 (ATCC  
CC CRL 8597).

XX Sequence 210 AA;

SQ Query Match 44.2%; Score 997; DB 21; Length 210;  
Best Local Similarity 85.2%; Pred. No. 2.1e-86;  
Matches 179; Conservative 16; Mismatches 11; Indels 4; Gaps 1;  
QY 110 MGGKWSKSSVVGWPTVYRMR-----RAEPAADGVGAASRDLEKHGAITSNTAATNAACA 165  
DB 1 mggkwsksrmggwsalrerrmrtaepaagdgavsdlekghgaitssntaataadca 60  
QY 166 WLEAQEEVEGFPVTPQVPLRPMYKAAVDLSHFLKEKGLEGLIHSQRQDILDLYIH 225  
DB 61 wleaqeeveegfpvtpqvlrmpmtykaaidishflkekgleglilwsqrqelldlyih 120  
QY 226 TQGYFPDMQNYTPGCVRYPLTFGWCYKLVPEPDKVEEANKGENTSLHLPVSLHGMDDP 285  
DB 121 tqgyfpdmqnytpgcvryplrtfgwciklvpepekeveeanegeenllhpmshgmeda 180  
QY 286 EREVLEWRFDSRLAFHHVARELHPEYFKNC 315  
DB 181 ekevlwrfdsklafhhmarelhpeyykdc 210

## RESULT 25

AAE04962  
ID AAE04962 standard; Protein; 217 AA.  
XX AC AAE04962;  
XX DT 10-SEP-2001 (first entry)  
XX DE HIV-1 Jrf1 Nef (G2A, LLAA) mutant protein.  
XX KW DNA vaccine; codon optimisation; Nef protein; HIV infection; anti-HIV;  
XX KW cell mediated immune response; cytotoxic T lymphocyte; CTL; mutein.  
XX KW human immunodeficiency virus-1; HIV-1; Jrf1 Nef; mutant; mutein.  
XX OS Human immunodeficiency virus type 1.  
XX OS Synthetic.  
XX FH Key Location/Qualifiers  
FT Misc-difference 2 /note= "Wild type myristylation site at Gly2 substituted  
FT with Ala"  
FT Misc-difference 174, 175 /note= "Wild type dileucine motif (Leu-Leu) substituted  
FT with Ala-Ala"

WO200143693-A2.

XX 21-JUN-2001.

XX 15-DEC-2000; 2000WO-US34162.

XX 17-DEC-1999; 99US-0172442.

XX (MERI ) MERCK & CO INC.

XX Shiver JW, Liang X, Fu T;

XX WPI; 2001-417878/44.

XX DR N-PSDB; AAD09605.

Novel DNA vaccine useful for inducing cell mediated immune response against virulent strains of HIV, comprises DNA expression vector, and DNA containing codon optimized open reading frame encoding Nef protein .

Claim 29; Page 18; 84pp; English.

The present invention relates to a DNA vaccine comprising a DNA expression vector, and a DNA molecule containing a codon optimised open reading frame encoding a Nef protein or its immunogenic derivative. The Nef protein or its derivative is expressed and generates an immune response which provides a substantial level of protection against HIV infection, upon administration of to a host. The DNA vaccine is useful for inducing a cell mediated immune (cytotoxic T lymphocyte, CTL) response against infection or disease caused by virulent strains of HIV, by administering the DNA vaccine into the tissue of the vertebrate host, preferably human. It is also useful for lowering transmission rate to previously uninfected individuals and/or for reducing levels of viral loads within an infected individual, so as to prolong the asymptomatic phase of HIV infection. The present sequence is human immunodeficiency virus-1 (HIV-1) Jrf1 Nef (G2A, LLAA) mutant protein. The different nef constructs of the invention are used in the exemplification of the invention.

Sequence 217 AA;

Query Match 44.2%; Score 996; DB 22; Length 217;  
Best Local Similarity 82.0%; Pred. No. 2.8e-86;  
Matches 178; Conservative 16; Mismatches 13; Indels 10; Gaps

QY 110 MGKSKSSVWGVTVRMRRAEPAAD-----GVGAASRDLEKHGAISSNTAA 159  
Db 1 mskgkskrsvpgwstvrsmrraeapadvrrtpeaavgvavsrdlekghaltsntaa 60

QY 160 TNAACAWLEAQEEVEEGFPVTPQVPLRPMTYKAAVDLSHLFLKEKGLEGLIHSORQDIL 219  
Db 61 tnadcaawleaqedeevgfvpqpvlrprmtkygavdlsnflkekggleglihsqrgdil 120

QY 220 DLWIYHTQGYPFDWQNYTGPCGYRVPYLTFGCWKLYPVEPKVPEEKANKGENTSLHPVSL 279  
Db 121 dlwyhtqqyfpdqwnytgpggrlftfgwcfklypvepekveeanegenncaahpmqs 180

QY 280 HGMDPEREVLEWRFSRLAFHHVARELHPFKNCT 316  
Db 181 hgiedpekevwlewrfsrlafhhvarelhpewykdc 217

RESULT 26  
AAR12262 ID AAR12262 standard; Protein; 210 AA.  
XX AC AAR12262;  
DT 20-AUG-1991 (first entry)  
XX XX  
XX HIV-1 strain OVI open reading frame (ORF) F protein.  
XX DE HIV-1; AIDS; retroviruses.  
XX KW Homo sapiens.  
OS OS  
PN US5019510-A.  
PP 28-OCT-1987; 87US-0113655.  
PD 28-MAY-1991.  
XX XX  
XX 28-OCT-1987; 87US-0113655.  
XX PR  
XX PA (INSP ) INST PASTEUR.  
PI Wain-Hobson S, Huet T, Delaporte E, Brun-Vezinet F;



```
QY 285 PEREYLFDSRLAFHHVARELHPEYKNC 315
Db 207 pekevwrfdfskiafhvharelhpeyykdc 237

RESULT 29
AAB69363
ID AAB69363 standard; Protein; 206 AA.
XX AAB69363;
AC
XX
XX
XX 20-APR-2001 (first entry)
XX
XX
XX HIV-1 non-subtype B clone 96ZM651-8 nef protein.
XX
XX HIV-1; human immunodeficiency virus; non-subtype B; gag; pol; env;
KW vpr; vif; vpr; tat; rev; nef; vaccine.
XX
XX Human immunodeficiency virus type 1.
OS
XX
XX WO200026416-A1.
XX
XX 11-MAY-2000.
XX
XX 25-OCT-1999; 99WO-US24837.
XX
XX 02-NOV-1998; 98US-0184418.
XX
XX (UABR-) UAB RES FOUND.
XX
XX Hahn BH, Shaw GM, Gao F;
PI
XX
XX WPI; 2000-365651/31.
XX
XX Novel genomic nucleic acids of non-subtype B human immunodeficiency
PT virus type 1 useful for detecting and treating AIDS comprises a
PT specific nucleotide sequence
XX
XX Claim 41; Fig 22; 131pp; English.
XX

The present invention provides the protein and coding sequences for a
number of human immunodeficiency virus (HIV) type 1 non-subtype B
isolates. The sequences shown include the near full-length coding
sequences from each isolate, and the env, pol, vif, vpr, vpu, gag, tat,
rev and nef proteins. These can be used to detect the presence of HIV-1
in a sample and to produce antibodies against non-subtype B HIV-1 virus.
These antibodies can be used in vaccines to prevent and treat HIV
infection.
XX
XX Sequence 206 AA;

Query Match 42.2%; Score 952; DB 21; Length 206;
Best Local Similarity 81.1%; Pred. No. 3.9e-82;
Matches 167; Conservative 24; Mismatches 15; Indels 0; Gaps 0;

QY 110 MGGKWSKSVGWGPIVRRMRRAEPAADGVGAASRDLEKHGAITSNTAATNAACAWLEA 169
Db 1 mdgkwsksivgwpavrrirrtetpaagvgasqldkkygaltssntttnaacawlea 60

QY 170 QEEEEVGFPVTPQVPLRPMYKAAVDLSHFLKEGGLIHSORRODILDLIWYHTQGY 229
Db 61 qeeeevgfvrpqvplrpmttykaavdlsfllkekgglegliyskrqrqeldlwyhtgqf 120

QY 230 FPDWQNYTPGPGVRYPLTFGCYKLVPEPDKVEANKGENTSLHPVSLHGMDDPEREV 289
Db 121 fpdwqnytpgpgvrypltfgcwcfklvpdpgeveeaneencllhpmssqgqgddhrev 180

QY 290 LEWRFDLSRLAFHHVARELHPEYKNC 315
Db 181 lkwfshlahkhmarelhpeyykdc 206

us-09-509-239-17.rag
```

```
RESULT 30
AAB86169
ID AAB86169 standard; Protein; 3025 AA.
XX
XX AAB86169;
AC
XX
XX 21-AUG-2001 (first entry)
XX
XX HIV-1 subtype C protein fragment #1.
XX
XX Infection; diagnosis; human; humoral immune response; antiviral;
KW cellular immune response; vaccine; treatment; gene therapy.
XX
XX Human immunodeficiency virus type 1.
OS
XX
XX Key Location/Qualifiers
FT Misc-difference 1..3025
FT /label= Xaa
FT /note= "Xaa represents a stop codon"
XX
XX DE10056747-A1.
XX
XX 31-MAY-2001.
XX
XX 16-NOV-2000; 2000DE-1056747.
XX
XX 16-NOV-1999; 99DE-1055089.
XX
XX (SHAO/) SHAO Y.
XX (GENE-) GENEART GMBH GES ANGEWANDTE BIOTECHNOLOG.
XX
XX Wagner R, Wolf H, Shao Y, Graf M;
PI
XX
XX WPI; 2001-336417/36.
XX N-PSDB; AAH20870.
XX
XX New nucleic acid sequences from a human immune deficiency virus
PT intersubtype, useful for treatment, prevention and diagnosis of
PT infection
XX
XX Disclosure: Fig 8A-O; 48pp; German.
XX
XX This invention describes a novel polynucleotide isolated from human
immunodeficiency virus type 1 subtype C/B' which can be used for the
induction of specific humoral and cellular immune responses (I) and
polypeptides (II) encoded by them, are useful in pharmaceuticals,
vaccines and diagnostic agents, particularly for treatment or prevention
of human immune deficiency virus-1 (HIV-1) infections, also for rational
design of test or therapeutic reagents, or gene therapy vectors.
XX Polypeptides, especially antibodies, specifically directed against (II)
are similarly useful as pharmaceutical and diagnostic agents. (I) are
specific for intersubtype C/B' of HIV-1 so are useful in regions
(particularly China and South-East Asia) where this subtype is prevalent.
XX The products of the invention have antiviral activity. This sequence
represents a protein encoded by the HIV-1 subtype C genome described in
the method of the invention.
XX
XX Sequence 3025 AA;

Query Match 42.1%; Score 948.5; DB 22; Length 3025;
Best Local Similarity 70.6%; Pred. No. 4.4e-80;
Matches 175; Conservative 22; Mismatches 42; Indels 9; Gaps 2;

QY 110 MGGKWSKSVGWGPIVRRMRRAEPAADGVGAASRDLEKHGAITSNTAATNAACAWLEA 169
Db 2723 mgkksksivgwpairrmrrrtetpaagvgavrdlekhgaltssntatnedcawlea 2782

QY 170 QEEEEVGFPVTPQVPLRPMYKAAVDLSHFLKEGGLIHSORRODILDLIWYHTQGY 229
Db 2783 qeeeevgfvrpqvplrpmttykaavdlsfllkekgglegliyskrqrqeldlwyhtggy 2842

QY 230 FPDWQNYTPGPGVRYPLTFGCYKLVPEPDKVEANKGENTSLHPVSLHGMDDPEREV 289
```

Db 2843 fpdwhnycpgpvrzpltfgrcfklvpvpreveeanegeedncllhpcqhgmeddhrev 2902  
Qy 290 LEWREDSRLAFHHVARELHPEYFKNCTSEPYDPR-----LEPWKHGSGQPKTACTNCY 342  
Db 2903 lkwkfdsqlahrarelhpefykdcxhrrdfprglstgafkevws--ggtgsgqpsnaa 2960  
Qy 343 CKKCCFHC 350  
Db 2961 ykqlifac 2968

Search completed: August 26, 2002, 08:12:27  
Job time: 236 sec

GenCore version 4.5  
Copyright (c) 1993 - 2000 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: August 26, 2002, 08:10:39 ; Search time 38.46 Seconds  
(without alignments)  
261.022 Million cell updates/sec

Title: US-09-509-239-17  
Perfect score: 2255  
Sequence: 1 CSHSSNMANTQMKSDKIII.....QSRGDTGPKETSGHHHHH 411

Scoring table:  
Gapop 10.0 , Gapext 0.5

Searched: 231628 seqs, 24425594 residues

Total number of hits satisfying chosen parameters: 231628

Minimum DB seq length: 0  
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

Database : Issued Patents-AA:  
1: /cgn2\_6/ptodata/2/1aa/5A-COMB.pep.\*  
2: /cgn2\_6/ptodata/2/1aa/5B-COMB.pep.\*  
3: /cgn2\_6/ptodata/2/1aa/6A-COMB.pep.\*  
4: /cgn2\_6/ptodata/2/1aa/6B-COMB.pep.\*  
5: /cgn2\_6/ptodata/2/1aa/PCTUS-COMB.pep.\*  
6: /cgn2\_6/ptodata/2/1aa/backfiles1.pep.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Query	Score	Match	Length	DB ID	Description
1	1138	50.5	239	4	US-08-679-493A-76	Sequence 76, Appl
2	1115	49.4	206	3	US-08-388-353-644	Sequence 644, App
3	1115	49.4	206	3	US-08-488-551B-644	Sequence 644, App
4	971	43.1	206	4	US-08-679-493A-75	Sequence 75, Appl
5	638	28.3	123	4	US-09-124-900-10	Sequence 10, Appl
6	585.5	26.0	278	4	US-09-485-885-21	Sequence 21, Appl
7	579	25.7	383	4	US-09-485-885-23	Sequence 23, Appl
8	576	25.5	220	4	US-09-485-885-8	Sequence 8, Appl
9	575	25.5	220	4	US-09-485-885-1	Sequence 1, Appl
10	572	25.4	227	4	US-09-485-885-16	Sequence 16, Appl
11	572	25.4	371	4	US-09-485-885-6	Sequence 6, Appl
12	569	25.2	227	4	US-09-485-885-19	Sequence 19, Appl
13	564	25.0	273	4	US-09-485-885-4	Sequence 4, Appl
14	498.5	22.1	151	4	US-08-679-493A-74	Sequence 74, Appl
15	486	21.6	86	1	US-08-450-257-1	Sequence 1, Appl
16	486	21.6	86	1	US-08-450-246-1	Sequence 1, Appl
17	486	21.6	86	1	US-08-450-098-1	Sequence 1, Appl
18	486	21.6	86	1	US-08-451-233-1	Sequence 1, Appl
19	486	21.6	86	1	US-08-450-236-1	Sequence 1, Appl
20	486	21.6	86	3	US-07-808-452-1	Sequence 1, Appl
21	486	21.6	86	3	US-09-030-613-19	Sequence 19, Appl
22	486	21.6	86	4	US-09-124-900-6	Sequence 6, Appl
23	486	21.6	86	4	US-09-451-905-19	Sequence 19, Appl
24	486	21.6	86	4	US-08-235-403-1	Sequence 1, Appl
25	486	21.6	86	5	PCT-US92-10770-1	Sequence 1, Appl
26	486	21.6	86	5	PCT-US95-06077-2	Sequence 2, Appl
27	478.5	21.2	266	3	US-08-815-809-5	Sequence 5, Appl

28	478.5	21.2	280	2	US-08-816-155B-43	Sequence 43, Appl
29	478.5	21.2	280	4	US-09-079-587-43	Sequence 43, Appl
30	470	20.8	86	2	US-08-505-210-1	Sequence 1, Appl
31	470	20.8	86	4	US-09-099-333-1	Sequence 1, Appl
32	466	20.7	83	2	US-08-417-210A-96	Sequence 96, Appl
33	455	20.2	82	1	US-08-053-079A-15	Sequence 15, Appl
34	439	19.5	253	2	US-08-659-251-4	Sequence 4, Appl
35	439	19.5	253	4	US-09-256-490-4	Sequence 4, Appl
36	439	19.5	253	5	PCT-US96-11445-4	Sequence 4, Appl
37	409	18.1	72	3	US-09-030-613-17	Sequence 17, Appl
38	409	18.1	72	4	US-09-451-905-17	Sequence 17, Appl
39	385	17.1	72	2	US-08-893-853-1	Sequence 1, Appl
40	385	17.1	72	4	US-09-113-921-1	Sequence 1, Appl
41	372	16.5	122	6	5304466-3	Patent No. 5304466
42	365.5	16.2	312	1	US-08-094-128A-27	Sequence 27, Appl
43	365.5	16.2	312	1	US-08-455-674-27	Sequence 27, Appl
44	365.5	16.2	312	1	US-08-455-992-27	Sequence 27, Appl
45	365.5	16.2	312	1	US-08-455-972-27	Sequence 27, Appl

ALIGNMENTS

RESULT 1  
US-08-679-493A-76  
; Sequence 76, Application US/08679493A  
; Patent No. 6303295  
; GENERAL INFORMATION:  
; APPLICANT: Taylor, Ethan W.  
; TITLE OF INVENTION: SELENOPROTEINS, CODING SEQUENCES AND METHODS  
; FILE REFERENCE: 55-95  
; CURRENT APPLICATION NUMBER: US/08/679,493A  
; PRIOR FILING DATE: 1996-07-12  
; PRIOR APPLICATION NUMBER: 60/001203  
; PRIOR FILING DATE: 1995-07-14  
; PRIOR APPLICATION NUMBER: 60/003,112  
; PRIOR FILING DATE: 1995-09-01  
; NUMBER OF SEQ ID NOS: 216  
; SOFTWARE: Patent In Ver. 2.0  
; SEQ ID NO 76  
; LENGTH: 239  
; TYPE: PRT  
; ORGANISM: Human immunodeficiency virus type 1  
; FEATURE:  
; NAME/KEY: VARIANT  
; LOCATION: (1)..(239)  
; OTHER INFORMATION: X os selenocysteine.  
US-08-679-493A-76

Query Match	50.5%	Score 1138;	DB 4;	Length 239;
Best Local Similarity	88.7%	Pred. No. 9.6e-107;		
Matches 211;	Conservative 2;	Mismatches 25;	Indels 0;	Gaps 0;
QY 110	MGKSKSKSVGWPTVVRMRRAEPAADGVGAASRDLEKHGAITSNTAATNACAWLEA 169			
Db 1	MGKSKSKSVGWPTVVRMRRAEPAADGVGAASRDLEKHGAITSNTAATNACAWLEA 60			
QY 170	QEEVEGVFPVTPQVPLRPMYTKAAVDLSHFLKEKGLEGILHSORQDILDLYHTQGY 229			
Db 61	QEEVEGVFPVTPQVPLRPMYTKAAVDLSHFLKEKGLEGILHSORQDILDLYHTQGY 120			
QY 230	FPDQWQNTPGGVRYPLTFGWCYKLVPEPKVVEANKGENTSLHPVSLHGMDPEREV 289			
Db 121	FPDQWQNTPGGVRYPLTFGWCYKLVPEPKVVEANKGENTSLHPVSLHGMDPEREV 180			
QY 290	LEWRDSDRLAFHHVARELHPEYFKNCTSEPVDPRLPEWHPGSOPTACTNCKYKCC 347			
Db 181	LEWRDSDRLAFHHVARELHPEYFKNCTSEPVDPRLPEWHPGSOPTACTNCKYKCC 238			

RESULT 2  
US-08-388-353-644

Sequence 644, Application US/08388353  
Patent No. 6010895  
GENERAL INFORMATION:  
APPLICANT: Deacon, Nicholas J.  
APPLICANT: Learmont, Jennifer C.  
APPLICANT: McPhee, Dale A.  
APPLICANT: Crowe, Suzanne  
APPLICANT: Cooper, David  
TITLE OF INVENTION: NON-PATHOGENIC STRAINS OF HIV-1  
NUMBER OF SEQUENCES: 800  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: Scully, Scott, Murphy & Presser  
STREET: 400 Garden City Plaza  
CITY: Garden City  
STATE: New York  
COUNTRY: United States  
ZIP: 11530  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: Patentin Release #1.0, Version #1.25  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/388,353  
FILING DATE: 14-FEB-1995  
CLASSIFICATION: 424  
ATTORNEY/AGENT INFORMATION:  
NAME: Digiglio, Frank S.  
REGISTRATION NUMBER: 31,346  
REFERENCE/DOCKET NUMBER: 9606  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: (516) 742-4343  
TELEFAX: (516) 742-4366  
TELEX: 230 901 SANS UR  
INFORMATION FOR SEQ ID NO: 644:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 206 amino acids  
TYPE: amino acid  
STRANDEDNESS: single  
TOPOLOGY: linear  
MOLECULE TYPE: protein  
US-08-388-353-644

Query Match 49.4%; Score 1115; DB 3; Length 206;  
Best Local Similarity 98.1%; Pred. No. 1.6e-104;  
Matches 202; Conservative 1; Mismatches 3; Indels 0; Gaps 0;  
Qy 110 MGGKWSKSSVVGWPTVRRMRRAEPAADGVGAASRDLEKHGAITSNTAATNAACAWLEA 169  
Db 1 MGGKWSKSSVVGWPAVRMRRAEPAADGVGAASRDLEKHGAITSNTAANNAACAWLEA 60  
Qy 170 QEEVEGFPVTPQVPLRPMTYKAAVDLSHFLKEKGGLEGLHSQRRQDILDWIYHTQGY 229  
Db 61 QEEVEGFPVTPQVPLRPMTYKAAVDLSHFLKEKGGLEGLHSQRRQDILDWIYHTQGY 120  
Qy 230 FPDQNYTPGCVRYPLTFGWCYKLVPEPKVPEKVEANKGENTSLHHPVSLHGMDDPEREV 289  
Db 121 FPDQNYTPGCVRYPLTFGWCYKLVPEPKVPEKVEANKGENTSLHHPVSLHGMDDPEREV 180  
Qy 290 LEWRDLSRLAFHHVARELHPEYFKNC 315  
Db 181 LEWRDLSRLAFHHVARELHPEYFKNC 206

RESULT 3  
US-08-488-551B-644  
Sequence 644, Application US/08488551B  
Patent No. 6015661  
GENERAL INFORMATION:  
APPLICANT: Nicholas J. Deacon  
APPLICANT: Dale A. McPhee  
APPLICANT: David Cooper

TITLE OF INVENTION: NON-PATHOGENIC STRAINS OF HIV-1  
NUMBER OF SEQUENCES: 841  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: SCULLY, SCOTT, MURPHY & PRESSER  
STREET: 400 GARDEN CITY PLAZA  
CITY: GARDEN CITY  
STATE: NEW YORK  
COUNTRY: U.S.A.  
ZIP: 11530-0299  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: Patentin Release #1.0, Version #1.25  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/488,551B  
FILING DATE: 07-JUN-1995  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: PM3864 (AU)  
FILING DATE: 14-FEB-1994  
APPLICATION NUMBER: PM4002 (AU)  
FILING DATE: 21-FEB-1994  
APPLICATION NUMBER: PNO284 (AU)  
FILING DATE: 23-DEC-1994  
APPLICATION NUMBER: US 08/388,353  
FILING DATE: 14-FEB-1995  
APPLICATION NUMBER: PNO3021/95  
FILING DATE: 17-MAY-1995  
ATTORNEY/AGENT INFORMATION:  
NAME: FRANK S. DIGIGLIO  
REFERENCE/DOCKET NUMBER: 96062  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: (516) 742-4343  
TELEFAX: (516) 742-4366  
INFORMATION FOR SEQ ID NO: 644:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 206 amino acids  
TYPE: amino acid  
STRANDEDNESS: single  
TOPOLOGY: linear  
MOLECULE TYPE: protein  
US-08-488-551B-644

Query Match 49.4%; Score 1115; DB 3; Length 206;  
Best Local Similarity 98.1%; Pred. No. 1.6e-104;  
Matches 202; Conservative 1; Mismatches 3; Indels 0; Gaps 0;  
Qy 110 MGGKWSKSSVVGWPTVRRMRRAEPAADGVGAASRDLEKHGAITSNTAATNAACAWLEA 169  
Db 1 MGGKWSKSSVVGWPAVRMRRAEPAADGVGAASRDLEKHGAITSNTAANNAACAWLEA 60  
Qy 170 QEEVEGFPVTPQVPLRPMTYKAAVDLSHFLKEKGGLEGLHSQRRQDILDWIYHTQGY 229  
Db 61 QEEVEGFPVTPQVPLRPMTYKAAVDLSHFLKEKGGLEGLHSQRRQDILDWIYHTQGY 120  
Qy 230 FPDQNYTPGCVRYPLTFGWCYKLVPEPKVPEKVEANKGENTSLHHPVSLHGMDDPEREV 289  
Db 121 FPDQNYTPGCVRYPLTFGWCYKLVPEPKVPEKVEANKGENTSLHHPVSLHGMDDPEREV 180  
Qy 290 LEWRDLSRLAFHHVARELHPEYFKNC 315  
Db 181 LEWRDLSRLAFHHVARELHPEYFKNC 206  
RESULT 4  
US-08-679-493A-75  
Sequence 75, Application US/08679493A  
Patent No. 6303295  
GENERAL INFORMATION:  
APPLICANT: Taylor, Ethan W.  
TITLE OF INVENTION: SELENOPROTEINS, CODING SEQUENCES AND METHODS  
FILE REFERENCE: 55-95

;; CURRENT APPLICATION NUMBER: US/08/679,493A  
;; CURRENT FILING DATE: 1996-07-12  
;; PRIOR APPLICATION NUMBER: 60/001203  
;; PRIOR FILING DATE: 1993-07-14  
;; PRIOR APPLICATION NUMBER: 60/003,112  
;; PRIOR FILING DATE: 1995-09-01  
;; NUMBER OF SEQ ID NOS: 216  
;; SOFTWARE: Patent In Ver. 2.0  
;; SEQ ID NO 75  
;; LENGTH: 206  
;; TYPE: PRT  
;; ORGANISM: Human immunodeficiency virus type 1  
;; NAME/KEY: VARIANT  
;; LOCATION: (1)..(206)  
;; OTHER INFORMATION: X is selenocysteine.  
US-08-679-493A-75

Query Match 43.1%; Score 971; DB 4; Length 206;  
Best Local Similarity 89.0%; Pred. No. 5.3e-90;  
Matches 186; Conservative 2; Mismatches 15; Indels 6; Gaps 3;  
QY 110 MGKWSKSSVVGWPTVRMRRAEPAADGVGAASRDLEKHGAISSNTAATNAACAWLEA 169  
Db 1 MGKWSKSSVVGWPTVRMRRAEPAADGVGAASRDLEKHGAISSNTAATNAACAWLEA 60  
QY 170 QEEVEGFPVTPQVPLRPMTYKAADVLSHFLKEKGLIHSQRDILDLYIHTQGY 226  
Db 61 QEEVEGFPVTPQVPLRPMTYKAADVLSHFLKEKGLIHSQRDILDLYIHTQGY 120  
QY 227 QGFPDQWNTPGVRYPLTFGWCYKLVPEPDKEEANKGENTSLLLHPVSLHGMDPE 286  
Db 119 R-LLPDQWNTPGVRYPLTFGWCYKLVPEPDKEEANKGENTSLLLHPVSLHGMDPE 177  
QY 287 REVLENRFDLSRLAFHVAHELHPEYKNC 315  
Db 178 REVLENRFDLSRLAFHVAHELHPEYKNC 206

RESULT 5  
US-09-124-900-10  
;; Sequence 10, Application US/09124900  
;; Patent No. 6268484  
;; GENERAL INFORMATION:  
;; APPLICANT: KATINGER, Hermann  
;; APPLICANT: BUCHACHER, Andrea  
;; APPLICANT: ERNST, Wolfgang  
;; APPLICANT: BALLAUN, Claudia  
;; APPLICANT: FURTSCHER, Martin  
;; APPLICANT: TRKOLA, Alexandra  
;; APPLICANT: PREDL, Renate  
;; APPLICANT: SCHMATZ, Christine  
;; APPLICANT: KLIMA, Annelies  
;; APPLICANT: STEINDL, Franz  
;; APPLICANT: MÜSTER, Thomas  
;; TITLE OF INVENTION: HIV-Vaccines  
;; FILE REFERENCE: 1939-112P  
;; CURRENT APPLICATION NUMBER: US/09/124,900  
;; CURRENT FILING DATE: 1998-07-30  
;; PRIOR APPLICATION NUMBER: PCT/EP95/01481  
;; PRIOR FILING DATE: 1995-04-19  
;; NUMBER OF SEQ ID NOS: 11  
;; SOFTWARE: Patent In version 3.0  
;; SEQ ID NO 10  
;; LENGTH: 123  
;; TYPE: PRT  
;; ORGANISM: Human immunodeficiency virus type 1  
US-09-124-900-10

Query Match 28.3%; Score 638; DB 4; Length 123;  
Best Local Similarity 97.6%; Pred. No. 9.3e-57;

Matches 120; Conservative 0; Mismatches 3; Indels 0; Gaps 0;  
QY 110 MGKWSKSSVVGWPTVRMRRAEPAADGVGAASRDLEKHGAISSNTAATNAACAWLEA 169  
Db 1 MGKWSKSSVVGWPAVVRMRRAEPAADGVGAASRDLEKHGAISSNTAANADCAWLEA 60  
QY 170 QEEVEGFPVTPQVPLRPMTYKAADVLSHFLKEKGLIHSQRDILDLYIHTQGY 229  
Db 61 QEEVEGFPVTPQVPLRPMTYKAADVLSHFLKEKGLIHSQRDILDLYIHTQGY 120  
QY 230 FPD 232  
Db 121 FPD 123  
RESULT 6  
US-09-485-885-21  
;; Sequence 21, Application US/09485885  
;; Patent No. 6342224  
;; GENERAL INFORMATION:  
;; APPLICANT: Bruck, Claudine  
;; APPLICANT: Cabezon Silva, Teresa  
;; APPLICANT: Delisse, Anne-Marie Eva Fernande  
;; APPLICANT: Gerard, Catherine Marie Ghislaine  
;; APPLICANT: Lombardo-Bencheikh, Angela  
;; TITLE OF INVENTION: Vaccine  
;; FILE REFERENCE: B45107  
;; CURRENT APPLICATION NUMBER: US/09/485,885  
;; CURRENT FILING DATE: 2000-02-18  
;; PRIOR APPLICATION NUMBER: PCT/EP98/05285  
;; PRIOR FILING DATE: 1998-08-17  
;; PRIOR APPLICATION NUMBER: GB 9717953.5  
;; PRIOR FILING DATE: 1997-08-22  
;; NUMBER OF SEQ ID NOS: 23  
;; SOFTWARE: FastSeq for Windows Version 3.0  
;; SEQ ID NO 21  
;; LENGTH: 278  
;; TYPE: PRT  
;; ORGANISM: Homo sapien  
US-09-485-885-21

Query Match 26.0%; Score 585.5; DB 4; Length 278;  
Best Local Similarity 36.0%; Pred. No. 6e-51;  
Matches 155; Conservative 26; Mismatches 73; Indels 177; Gaps 12;  
QY 2 SSHSSNNANTQMSDKIIIAHARGASGYLPHTLESKALAFQAQADYLEQDLAMTKDGLRV 61  
Db 4 SSHSSNNANTQMSDKIIIAHARGASGYLPHTLESKALAFQAQADYLEQDLAMTKDGLRV 63  
QY 62 VIHDFLDGLTDVAKFPFHRKDGRIYVDFTLKEIQSLEMTENFETMGKWSKSSVWG 121  
Db 64 VIHDFLDGLTDVAKFPFHRKDGRIYVDFTLKEIQSLEMTENFETMA----- 113  
QY 122 WPTVRMRRAEPAADGVGAASRDLEKHGAISSNTAATN--AACAW-----LEAQEEVEV 175  
Db 114 -----REFDP-----TRRYPKLPDLCTELNTSLQDIEITCYCKVLTETVEFEF 158  
QY 176 GFPVTPQVPLRPMTYKAADVLSHFLKEKGLIHSQRDILDLYIHTQGYFPDQW 235  
Db 159 AFK-----DLFVVY----- 167  
QY 236 YTPGPGVRYPLTFGWCYKLV-----PVEPDKEE-ANKGENTSLLLHPVSLH 280  
Db 168 -----RDSIPHAACHKIDFYSRIRELRHYSVYDGTLEKLTNTGLYNLLIRLCRCQ 220  
QY 281 GMDPPERVLEWRDLSRLAFHVAHELHPEYKNCISEPVDPRLEKHPGSKPTACTN 340  
Db 221 KPLNPAEKLK--HLNEKRREPHNIAGHYRGO----- 248  
QY 341 CYCKKCCFHCQVCFITKALGISYGRKRORRRPPGSGSTHQSLSKQPTSQSRGDDTGP 400  
Db 249 --CHSC-----NRAQERLORRRETQ----- 268



QY 401 KETSGHHHHH 411  
Db 269 -VTSGHHHHH 278

RESULT 7  
US-09-485-885-23  
; Sequence 23, Application US/09485885  
; Patent No. 6342224  
; GENERAL INFORMATION:  
; APPLICANT: Bruck, Claudine  
; APPLICANT: Cabezon Silva, Teresa  
; APPLICANT: Delisse, Anne-Marie Eva Fernande  
; APPLICANT: Gerard, Catherine Marie Ghislaine  
; APPLICANT: Lombardo-Bencheikh, Angela  
; TITLE OF INVENTION: Vaccine  
; FILE REFERENCE: B45107  
; CURRENT APPLICATION NUMBER: US/09/485,885  
; CURRENT FILING DATE: 2000-02-18  
; PRIOR APPLICATION NUMBER: PCT/EP98/05285  
; PRIOR FILING DATE: 1998-08-17  
; PRIOR APPLICATION NUMBER: GB 9717953.5  
; PRIOR FILING DATE: 1997-08-22  
; NUMBER OF SEQ ID NOS: 23  
; SOFTWARE: FastSeq for Windows Version 3.0  
; SEQ ID NO 23  
; LENGTH: 383  
; TYPE: PRT  
; ORGANISM: Homo sapien  
US-09-485-885-23

Query Match 25.78; Score 579; DB 4; Length 383;  
Best Local Similarity 34.58; Pred. No. 4.4e-50;  
Matches 167; Conservative 21; Mismatches 118; Indels 178; Gaps 11;  
QY 2 SSSSNMANTOMKSKIIIAHRGASGYLPEHTLESKALAFQAQADYLDQDLAMTKDGRV 61  
Db 4 SSSSNMANTOMKSKIIIAHRGASGYLPEHTLESKALAFQAQADYLDQDLAMTKDGRV 63  
QY 62 VIHDFLDGLTDVAKKFPHRHRKDGYYVIDFTLKEIOSLEMTENFETMGG----- 112  
Db 64 VIHDFLDGLTDVAKKFPHRHRKDGYYVIDFTLKEIOSLEMTENFETMARFEDPTRPY 123  
QY 113 -----KWSKSSVVGWPTVRMRRAEPAADGVGAASRDLEKHGATIS----- 154  
Db 124 KLPLDCTELNTSLQDIEITCYCKTVLETFEFAFKDLFVYVRDSDIPHAACHKCIDFY 183  
QY 155 -----SNTAATNAACAWLEAQEEEEVEGFPVTPQVPLRPMYKAA 193  
Db 184 SPIRELHYSDVGYDTLEKLTNTGLYLLRCLRCQK-----PLNPAEKLRLHNEK-- 235  
QY 194 VDLSHFLKEKGLELIHS---ORQDILDWIYHTOGYFFPDWQNYTPGPGVRYPLTFGW 250  
Db 236 ---RRFHNIAHYRQCHSCCNRAQ----- 258  
QY 251 CYKLIVPEPKVEEANKGENTSLH-----PVSL---HGMDPPRE 288  
Db 259 -----ERLQRRRETQVMHGPKATLQDIVLHLEPONEIPVDLLCHEQLSDSEE 306  
QY 289 VLEWRDRLAFPHVARELHPE-YFKNCTSEPVDRLEPWRKPGSQPKTACTNCKKCC 347  
Db 307 NDE--IDGVNHOHLPARAEQPHRTMLCMCKCEARIELVWESSADDLRAQOQLFNLTS 364  
QY 348 FHCQVCFTIKALGISYGRKKRRRRPQGSQTHQVSLSKQPTOSRQDPTGPKETSGHH 407  
Db 365 FVCPWC-----ASQOTSQHH 379  
QY 408 HHHH 411  
Db 380 HHHH 383

RESULT 8  
US-09-485-885-8  
; Sequence 8, Application US/09485885  
; Patent No. 6342224  
; GENERAL INFORMATION:  
; APPLICANT: Bruck, Claudine  
; APPLICANT: Cabezon Silva, Teresa  
; APPLICANT: Delisse, Anne-Marie Eva Fernande  
; APPLICANT: Gerard, Catherine Marie Ghislaine  
; APPLICANT: Lombardo-Bencheikh, Angela  
; TITLE OF INVENTION: Vaccine  
; FILE REFERENCE: B45107  
; CURRENT APPLICATION NUMBER: US/09/485,885  
; CURRENT FILING DATE: 2000-02-18  
; PRIOR APPLICATION NUMBER: PCT/EP98/05285  
; PRIOR FILING DATE: 1998-08-17  
; PRIOR APPLICATION NUMBER: GB 9717953.5  
; PRIOR FILING DATE: 1997-08-22  
; NUMBER OF SEQ ID NOS: 23  
; SOFTWARE: FastSeq for Windows Version 3.0  
; SEQ ID NO 8  
; LENGTH: 220  
; TYPE: PRT  
; ORGANISM: Homo sapien  
US-09-485-885-8

Query Match 25.58; Score 576; DB 4; Length 220;  
Best Local Similarity 66.58; Pred. No. 3.9e-50;  
Matches 123; Conservative 9; Mismatches 27; Indels 26; Gaps 3;  
QY 2 SSSSNMANTOMKSKIIIAHRGASGYLPEHTLESKALAFQAQADYLDQDLAMTKDGRV 61  
Db 4 SSSSNMANTOMKSKIIIAHRGASGYLPEHTLESKALAFQAQADYLDQDLAMTKDGRV 63  
QY 62 VIHDFLDGLTDVAKKFPHRHRKDGYYVIDFTLKEIOSLEMTENFETMGGKWSKSSVVG 121  
Db 64 VIHDFLDGLTDVAKKFPHRHRKDGYYVIDFTLKEIOSLEMTENFETMAMHGD----- 118  
QY 122 WPTVRMRRAEPAADGVGAASRDLEKHGATISNTAATNAACAWLEAQEEEEVEGFPVTP 181  
Db 119 -PTLHMYMLDLP-----ETDLYGYQLNDS-----SEEDIDGPAGQ 157  
QY 182 QVPLR 186  
Db 158 AEPDR 162

RESULT 9  
US-09-485-885-1  
; Sequence 1, Application US/09485885  
; Patent No. 6342224  
; GENERAL INFORMATION:  
; APPLICANT: Bruck, Claudine  
; APPLICANT: Cabezon Silva, Teresa  
; APPLICANT: Delisse, Anne-Marie Eva Fernande  
; APPLICANT: Gerard, Catherine Marie Ghislaine  
; APPLICANT: Lombardo-Bencheikh, Angela  
; TITLE OF INVENTION: Vaccine  
; FILE REFERENCE: B45107  
; CURRENT APPLICATION NUMBER: US/09/485,885  
; CURRENT FILING DATE: 2000-02-18  
; PRIOR APPLICATION NUMBER: PCT/EP98/05285  
; PRIOR FILING DATE: 1998-08-17  
; PRIOR APPLICATION NUMBER: GB 9717953.5  
; PRIOR FILING DATE: 1997-08-22  
; NUMBER OF SEQ ID NOS: 23  
; SOFTWARE: FastSeq for Windows Version 3.0  
; SEQ ID NO 1  
; LENGTH: 220  
; TYPE: PRT  
; ORGANISM: Homo sapien

US-09-485-885-1

Query Match 25.5%; Score 575; DB 4; Length 220;  
Best Local Similarity 66.5%; Pred. No. 4.9e-50;  
Matches 123; Conservative 9; Mismatches 27; Indels 26; Gaps 3;

QY 2 SSHSNMANTOMKSDKIIIAHRCAGSYLPEHTLESKALAFQAQADYLEODLMTKGRV 61  
DB 4 SSHSNMANTOMKSDKIIIAHRCAGSYLPEHTLESKALAFQAQADYLEODLMTKGRV 63

QY 62 VIHDFDLGLTDVAKKFPFHRKRGYVYIDFTLKEIQSLEMTENFETMGKWSKSSVVG 121  
DB 64 VIHDFDLGLTDVAKKFPFHRKRGYVYIDFTLKEIQSLEMTENFETMAHGDGT----- 118

QY 122 WPTVRMRRAEPAADGVGAASRDLEKKGAISSNTAATNAACAWLEAOEEVEEVPVTP 181  
DB 119 -PTLHEMLDLP-----ETDLYCYELNDS-----SDEEDIDGPAG 157

QY 182 QVPLR 186  
DB 158 AEPDR 162

## RESULT 10

US-09-485-885-16  
; Sequence 16, Application US/09485885  
; Patent No. 6342224  
; GENERAL INFORMATION:  
; APPLICANT: Bruck, Claudine  
; APPLICANT: Cabazon Silva, Teresa  
; APPLICANT: Delisse, Anne-Marie Eva Fernande  
; APPLICANT: Gerard, Catherine Marie Ghislaine  
; APPLICANT: Lombardo-Bencheikh, Angela  
; TITLE OF INVENTION: Vaccine  
; FILE REFERENCE: B45107  
; CURRENT APPLICATION NUMBER: US/09/485,885  
; CURRENT FILING DATE: 2000-02-18  
; PRIOR APPLICATION NUMBER: PCT/EP98/05285  
; PRIOR FILING DATE: 1998-08-17  
; PRIOR APPLICATION NUMBER: GB 9717953.5  
; PRIOR FILING DATE: 1997-08-22  
; NUMBER OF SEQ ID NOS: 23  
; SOFTWARE: FastSeq for Windows Version 3.0  
; SEQ ID NO 16  
; LENGTH: 227  
; TYPE: PRT  
; ORGANISM: Homo sapien  
US-09-485-885-16

Query Match 25.4%; Score 572; DB 4; Length 227;  
Best Local Similarity 36.5%; Pred. No. 1e-49;  
Matches 150; Conservative 14; Mismatches 59; Indels 188; Gaps 10;

QY 2 SSHSNMANTOMKSDKIIIAHRCAGSYLPEHTLESKALAFQAQADYLEODLMTKGRV 61  
DB 4 SSHSNMANTOMKSDKIIIAHRCAGSYLPEHTLESKALAFQAQADYLEODLMTKGRV 63

QY 62 VIHDFDLGLTDVAKKFPFHRKRGYVYIDFTLKEIQSLEMTENFETMGKWSKSSVVG 121  
DB 64 VIHDFDLGLTDVAKKFPFHRKRGYVYIDFTLKEIQSLEMTENFETM----- 112

QY 122 WPTVRMRRAEPAADGVGAASRDLEKKGAISSNTAATNAACAWLEAOEEVEEVPVTP 181  
DB 113 -----AMHGPKATLQDIVLH-----LEPONE----- 133

QY 182 QVPLRPTYKAAVDLSHFLKEKGLSLSORRQDILDLIWYHTOGYFDPQWNTYTPGC 241  
DB 134 -IP-----VDL-----LCHEQLSDS----- 147

QY 242 VRYPLTFGWCYKLVPEPKVEEANKGENTSLHHPVSLHGMDDPEREVLWRPDSRLAPH 301  
| : : : : |

DB 148 -----EENDEIDEVN-----HQ-----H 161

QY 302 HVARELHPE-YFNKCTSEPVDPRLPEPKHPSQPKTACTNCKKCFHCQVCFITKALG 360  
| : : : : |

DB 162 LPARAEAPORHTMLCMCKCEARIELWESSADDLRAFOQLFNTLSFVCPWC----- 214

QY 361 ISVGRKKRRORRRPPGOSQTHVSLSKQPTSQSRGDPGPKETSGHHHHH 411  
: : : : :  
DB 215 -----ASQOTSCHHHHH 227

RESULT 11  
US-09-485-885-6  
; Sequence 6, Application US/09485885  
; Patent No. 6342224  
; GENERAL INFORMATION:  
; APPLICANT: Bruck, Claudine  
; APPLICANT: Cabazon Silva, Teresa  
; APPLICANT: Delisse, Anne-Marie Eva Fernande  
; APPLICANT: Gerard, Catherine Marie Ghislaine  
; APPLICANT: Lombardo-Bencheikh, Angela  
; TITLE OF INVENTION: Vaccine  
; FILE REFERENCE: B45107  
; CURRENT APPLICATION NUMBER: US/09/485,885  
; CURRENT FILING DATE: 2000-02-18  
; PRIOR APPLICATION NUMBER: PCT/EP98/05285  
; PRIOR FILING DATE: 1998-08-17  
; PRIOR APPLICATION NUMBER: GB 9717953.5  
; PRIOR FILING DATE: 1997-08-22  
; NUMBER OF SEQ ID NOS: 23  
; SOFTWARE: FastSeq for Windows Version 3.0  
; SEQ ID NO 6  
; LENGTH: 371  
; TYPE: PRT  
; ORGANISM: Homo sapien  
US-09-485-885-6

Query Match 25.4%; Score 572; DB 4; Length 371;  
Best Local Similarity 34.5%; Pred. No. 2.1e-49;  
Matches 164; Conservative 27; Mismatches 112; Indels 172; Gaps 14;

QY 2 SSHSNMANTOMKSDKIIIAHRCAGSYLPEHTLESKALAFQAQADYLEODLMTKGRV 61  
DB 4 SSHSNMANTOMKSDKIIIAHRCAGSYLPEHTLESKALAFQAQADYLEODLMTKGRV 63

QY 62 VIHDFDLGLTDVAKKFPFHRKRGYVYIDFTLKEIQSLEMTENFETMGKWSKSSVVG 121  
DB 64 VIHDFDLGLTDVAKKFPFHRKRGYVYIDFTLKEIQSLEMTENFETM-----AM 114

QY 122 WPTVRMRRAEPAADGVGAASRDLEKKGAISSNTAATNAACAWLEAOEEVEEVPVTP 181  
DB 115 FQDQERPRKLPOLCTELOTTIHI-----ILECVYCKQ----- 149

QY 182 QVPLRPTYKAAVDLSHFLKEKGLSLSORRQDILDLIWYHTOG-----YF 230  
| : : : : |

DB 150 --LLREVIDFA-----FRDLCIVIRGDNPNYAVCDKCLFY 183

QY 231 PQWNTYTPGVRGYPLTFGWCYKLVPEPKVEEANKGENTSLHHPVSLHGMDDPEREVL 290  
| : : : : |

DB 184 SKTSEYR-----HYCYSLYGTLE--QYKPKCLDLLIRINCOKPLCPKEK-- 228

QY 291 EWRFDRLAFHVAHELHPEYFNKCTSEVPDRLE-----PWKHP-----GSQPKTACTNC 341  
| : : : : |

DB 229 QRHLDDKORFHNIRGRWTCMCCSSRRTRRETQLMHGDTPTLHGYMLDLPETDLYC 288  
| : : : : |

QY 342 Y-----CKKCCFHCQVC-----FIT 356  
| : : : : |

DB 289 YEOLNDSSEEDIDGPAGOEPRAHYNIIVTFCCCKDCLRLCVQSTHVDFITLEDLM 348  
| : : : : |

QY 357 KALGISYGRKKRRQRPPGOSQTHVSLSKQPTSQSRGDPGPKETSGHHHHH 411  
| : : : : |

DB 349 GTLGIVC-----PICSQ-----KFTSGHHHHH 371

```
RESULT 12
US-09-485-885-19
; Sequence 19, Application US/09485885
; Patent No. 6342224
; GENERAL INFORMATION:
; APPLICANT: Bruck, Claudine
; APPLICANT: Cabazon Silva, Teresa
; APPLICANT: Delisse, Anne-Marie Eva Fernandez
; APPLICANT: Gerard, Catherine Marie Ghislaine
; APPLICANT: Lombardo-Bencheikh, Angela
; TITLE OF INVENTION: Vaccine
; FILE REFERENCE: B45107
; CURRENT APPLICATION NUMBER: US/09/485, 885
; CURRENT FILING DATE: 2000-02-18
; PRIOR APPLICATION NUMBER: PCT/EP98/05285
; PRIOR FILING DATE: 1998-08-17
; PRIOR APPLICATION NUMBER: GB 9717953.5
; PRIOR FILING DATE: 1997-08-22
; NUMBER OF SEQ ID NOS: 23
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 19
; LENGTH: 227
; TYPE: PRT
; ORGANISM: Homo sapien
US-09-485-885-19

Query Match      25.2%; Score 569; DB 4; Length 227;
Best Local Similarity 36.3%; Pred. No. 2,1e-49;
Matches 149; Conservative 15; Mismatches 59; Indels 188; Gaps 9;

QY 2 SSHSNMANTQKSDKIIIAHNGASGYLPENTLESKALAFQAQADYLEODLMTKDGRLV 61
DB 4 SSHSNMANTQKSDKIIIAHNGASGYLPENTLESKALAFQAQADYLEODLMTKDGRLV 63
QY 62 VINDHFLDGLTDVAKKFPRIHRRKDGRIYVIDFTLKEIQSLKTEMTENFTNGGKWSKVY 121
DB 64 VINDHFLDGLTDVAKKFPRIHRRKDGRIYVIDFTLKEIQSLKTEMTENFTNGGKWSKVY 112
QY 122 WPTVERMRARADGVCASRDLEKIGALITSSNTAATNACAMLEAEDEEVEGFRVPT 181
DB 113 -----AMNGPKATLDIVL-----LEPONE----- 133
QY 182 QVLPKMTYKAADVLSHFLKKGLEGILHSORODILDLMTYNTGVRPMQNTYTPRG 241
DB 134 -IP-----VDLL-----GHQQLSDSEENDEID-----G 156
QY 242 VRRPLTFCMCYKLVPRDKVEANKGENTSLIHVSLSHGMDREVELMWRDRLAHN 301
DB 157 VNRQ-----H 161
QY 302 HVARELHPE-YFKNCTSEVPDRLEPKHNGSOPKTACTNCKSKCCSFHOCVFTTKALG 360
DB 162 LPARRAEPORHTYLMCCSCSEARIELVYESSADLRAFOOLFNTLSFVCRMC----- 214
QY 361 ISYGRKKRRORRRPQSGTHOVSLSKOPTSOSRGDPTGPKETSGHNHNN 411
DB 215 -----ASQQTSGHNHNN 227

RESULT 13
US-09-485-885-4
; Sequence 4, Application US/09485885
; Patent No. 6342224
; GENERAL INFORMATION:
; APPLICANT: Bruck, Claudine
; APPLICANT: Cabazon Silva, Teresa
; APPLICANT: Delisse, Anne-Marie Eva Fernandez
; APPLICANT: Gerard, Catherine Marie Ghislaine
; APPLICANT: Lombardo-Bencheikh, Angela
; TITLE OF INVENTION: Vaccine

FILE REFERENCE: B45107
; CURRENT APPLICATION NUMBER: US/09/485, 885
; CURRENT FILING DATE: 2000-02-18
; PRIOR APPLICATION NUMBER: PCT/EP98/05285
; PRIOR FILING DATE: 1998-08-17
; PRIOR APPLICATION NUMBER: GB 9717953.5
; PRIOR FILING DATE: 1997-08-22
; NUMBER OF SEQ ID NOS: 23
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 4
; LENGTH: 273
; TYPE: PRT
; ORGANISM: Homo sapien
US-09-485-885-4

Query Match      25.0%; Score 564; DB 4; Length 273;
Best Local Similarity 35.9%; Pred. No. 8,6e-49;
Matches 151; Conservative 24; Mismatches 84; Indels 162; Gaps 12;

QY 2 SSHSNMANTQKSDKIIIAHNGASGYLPENTLESKALAFQAQADYLEODLMTKDGRLV 61
DB 4 SSHSNMANTQKSDKIIIAHNGASGYLPENTLESKALAFQAQADYLEODLMTKDGRLV 63
QY 62 VINDHFLDGLTDVAKKFPRIHRRKDGRIYVIDFTLKEIQSLKTEMTENFTNGGKWSKVY 121
DB 64 VINDHFLDGLTDVAKKFPRIHRRKDGRIYVIDFTLKEIQSLKTEMTENFTNGGKWSKVY 114
QY 122 WPTVERMRARADGVCASRDLEKIGALITSSNTAATNACAMLEAEDEEVEGFRVPT 181
DB 115 FDDPOERPKQLQTLTDTTHTI-----ILECVICKQO----- 149
QY 182 QVLPKMTYKAADVLSHFLKKGLEGILHSORODILDLMTYNTG-----YF 230
DB 150 --LREYVDF-----FRDLCTIVYRDGNPYAVCDCKLKFY 183
QY 231 PDMQNTYTPRGVRYPLTFCMCYKLVPRDKVEANKGENTSLIHVSLSHGMDREVEL 290
DB 184 SKISER-----HYCYSYLTGLE--QYKPKLDDLIRINQKPLSPREK-- 228
QY 291 EWRFDRLAHNVARELHPEYFKNCTSEVPDRLEPKHNGSOPKTACTNCKSKCCSFNC 350
DB 229 QNHLDKKORFHNIRG-----W-----TGCMSC----- 253
QY 351 QVCFITKALGISYGRKKRRORRRPQSGTHOVSLSKOPTSOSRGDPTGPKETSGHNHNN 410
DB 254 -----RSSRTRE-----TOLTSGHNHNN 272
QY 411 H 411
DB 273 H 273

RESULT 14
US-08-679-493A-74
; Sequence 74, Application US/08679493A
; Patent No. 6303295
; GENERAL INFORMATION:
; APPLICANT: Taylor, Ethan W.
; TITLE OF INVENTION: SELENOPROTEINS, CODING SEQUENCES AND METHODS
; FILE REFERENCE: 55-95
; CURRENT APPLICATION NUMBER: US/08/679,493A
; CURRENT FILING DATE: 1996-07-12
; PRIOR APPLICATION NUMBER: 60/001203
; PRIOR FILING DATE: 1995-07-14
; PRIOR APPLICATION NUMBER: 60/003,112
; PRIOR FILING DATE: 1995-09-01
; NUMBER OF SEQ ID NOS: 216
; SOFTWARE: Patent Ver. 2.0
; SEQ ID NO 74
; LENGTH: 151
; TYPE: PRT
; ORGANISM: Human immunodeficiency virus type 1
```

FEATURE:  
NAME/KEY: VARIANT  
LOCATION: (1) (151)  
OTHER INFORMATION: X is selenocysteine.  
US-08-679-493A-74

Query Match 22.1%; Score 498.5; DB 4; Length 151;  
Best Local Similarity 73.0%; Pred. No. 1.5e-42;  
Matches 103; Conservative 4; Mismatches 7; Indels 27; Gaps 3;  
QY 110 MGKWSKSSVWGVTVRMRRAEPAADGGAASRDLEKHGAITSSNTAATNAACAWLEA 169  
DB 1 MGKWSKSSVWGVTVRMRRAEPAADGGAASRDLEKHGAITSSNTAATNAACAWLEA 60  
QY 170 QEEVEGVFTVQVPLRPMPTTYAAVDLSHFLKEKGG----- 205  
DB 61 QEEVEGVFTVQVPLRPMPTTYAAVDLSHFLKEKGGTGRANSLPTKTRYPXSVLDLPHRL 120  
QY 206 LEGL--IHSORRODI-LDLWI 223  
DB 121 LPXLAELHTRARGGISTDLWM 141

RESULT 15

US-08-450-257-1  
Sequence 1, Application US/08450257  
Patent No. 5652122  
GENERAL INFORMATION:  
APPLICANT: FRANKEL, Alan  
APPLICANT: PABO, Carl  
APPLICANT: BARSOUM, James G.  
APPLICANT: FAWELL, Stephen E.  
APPLICANT: PEPINSKY, R. B.  
TITLE OF INVENTION: TAT-DERIVED TRANSPORT POLYPEPTIDES  
NUMBER OF SEQUENCES: 69  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: FISH & NEAVE  
STREET: 1251 Avenue of the Americas  
CITY: New York  
STATE: New York  
COUNTRY: USA  
ZIP: 10020  
COMPUTER READABLE FORM:  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: Patent In Release #1.0, Version #1.25  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/450,257  
FILING DATE: 25-MAY-1995  
CLASSIFICATION: 514  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: US 08/235,403  
FILING DATE: 28-APR-1994  
APPLICATION NUMBER: US 07/934,375  
FILING DATE: 21-AUG-1992  
APPLICATION NUMBER: US 07/098,766  
FILING DATE: 28-JUL-1993  
APPLICATION NUMBER: PCT/US93/07833  
FILING DATE: 19-AUG-1993  
APPLICATION NUMBER: US 07/454,450  
FILING DATE: 21-DEC-1989  
APPLICATION NUMBER: US 07/636,662  
FILING DATE: 02-JAN-1991  
APPLICATION NUMBER: US 08/158,015  
FILING DATE: 24-NOV-1993  
ATTORNEY/AGENT INFORMATION:  
NAME: Haley Jr., James F.  
REGISTRATION NUMBER: 27,794  
REFERENCE/DOCKET NUMBER: B170 CIP 2  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: (212) 596-9000

TELEFAX: (212) 596-9090  
TELEX: 14-8367  
INFORMATION FOR SEQ ID NO: 1:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 86 amino acids  
TYPE: amino acid  
TOPOLOGY: linear  
MOLSCULE TYPE: protein  
ORIGINAL SOURCE:  
ORGANISM: human immunodeficiency virus  
STRAIN: type 1  
US-08-450-257-1  
Query Match 21.6%; Score 486; DB 1; Length 86;  
Best Local Similarity 100.0%; Pred. No. 1.2e-41;  
Matches 85; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
QY 318 EPVDPRLPEWKHPGSPKTNCTCYCKKCFHCQVCFITKALGISYGRKKRRRRPPQG 377  
DB 2 EPVDPRLPEWKHPGSPKTNCTCYCKKCFHCQVCFITKALGISYGRKKRRRRPPQG 61  
QY 378 SOTHQVSLSKQPTSQSRGDPGPK 402  
DB 62 SOTHQVSLSKQPTSQSRGDPGPK 86

RESULT 16

US-08-450-246-1  
Sequence 1, Application US/08450246  
Patent No. 5670617  
GENERAL INFORMATION:  
APPLICANT: FRANKEL, Alan  
APPLICANT: PABO, Carl  
APPLICANT: BARSOUM, James G.  
APPLICANT: FAWELL, Stephen E.  
APPLICANT: PEPINSKY, R. B.  
TITLE OF INVENTION: TAT-DERIVED TRANSPORT POLYPEPTIDES  
NUMBER OF SEQUENCES: 69  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: FISH & NEAVE  
STREET: 1251 Avenue of the Americas  
CITY: New York  
STATE: New York  
COUNTRY: USA  
ZIP: 10020  
COMPUTER READABLE FORM:  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: Patent In Release #1.0, Version #1.25  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/450,246  
FILING DATE: 25-MAY-1995  
CLASSIFICATION: 514  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: US 08/235,403  
FILING DATE: 28-APR-1994  
APPLICATION NUMBER: US 07/934,375  
FILING DATE: 21-AUG-1992  
APPLICATION NUMBER: US 07/098,766  
FILING DATE: 28-JUL-1993  
APPLICATION NUMBER: PCT/US93/07833  
FILING DATE: 19-AUG-1993  
APPLICATION NUMBER: US 07/454,450  
FILING DATE: 21-DEC-1989  
APPLICATION NUMBER: US 07/636,662  
FILING DATE: 02-JAN-1991  
APPLICATION NUMBER: US 08/158,015  
FILING DATE: 24-NOV-1993  
ATTORNEY/AGENT INFORMATION:  
NAME: Haley Jr., James F.  
REGISTRATION NUMBER: 27,794

REFERENCE/DOCKET NUMBER: B170 CIP 2  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: (212) 596-9000  
TELEFAX: (212) 596-9090  
TELEX: 14-8367  
INFORMATION FOR SEQ ID NO: 1:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 86 amino acids  
TYPE: amino acid  
TOPOLOGY: linear  
MOLECULE TYPE: protein  
ORIGINAL SOURCE: human immunodeficiency virus  
STRAIN: type 1  
US-08-450-246-1

Query Match 21.6%; Score 486; DB 1; Length 86;  
Best Local Similarity 100.0%; Pred. No. 1.2e-41;  
Matches 85; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
QY 318 EPVDPRLPEPKHPSQPKTACTNCTCYCKKCCFHCQVCFITKALGISYGRKKRRRRPPQG 377  
DB 2 EPVDPRLPEPKHPSQPKTACTNCTCYCKKCCFHCQVCFITKALGISYGRKKRRRRPPQG 61  
QY 378 SOTHQVSLSKOPTSQSRGDPGPK 402  
DB 62 SOTHQVSLSKOPTSQSRGDPGPK 86

RESULT 17  
US-08-450-098-1  
Sequence 1, Application US/08450098  
Patent No. 5674980  
GENERAL INFORMATION:  
APPLICANT: FRANKEL, Alan  
APPLICANT: PABO, Carl  
APPLICANT: BARSOUM, James G.  
APPLICANT: FAWELL, Stephen E.  
APPLICANT: PEPINSKY, R. B.  
TITLE OF INVENTION: TAT-DERIVED TRANSPORT POLYPEPTIDES  
NUMBER OF SEQUENCES: 69  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: FISH & NEAVE  
STREET: 1251 Avenue of the Americas  
CITY: New York  
STATE: New York  
COUNTRY: USA  
ZIP: 10020  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: Patent In Release #1.0, Version #1.25  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/450.098  
FILING DATE: 25-MAY-1995  
CLASSIFICATION: 514  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: US 08/235,403  
FILING DATE: 28-APR-1994  
APPLICATION NUMBER: US 07/934,375  
FILING DATE: 21-AUG-1992  
APPLICATION NUMBER: US 07/098,766  
FILING DATE: 28-JUL-1993  
APPLICATION NUMBER: PCT/US93/07833  
FILING DATE: 19-AUG-1993  
APPLICATION NUMBER: US 07/454,450  
FILING DATE: 21-DEC-1989  
APPLICATION NUMBER: US 07/636,662  
FILING DATE: 02-JAN-1991  
APPLICATION NUMBER: US 08/158,015  
FILING DATE: 24-NOV-1993

ATTORNEY/AGENT INFORMATION:  
NAME: Haley Jr., James F.  
REGISTRATION NUMBER: 27,794  
REFERENCE/DOCKET NUMBER: B170 CIP 2  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: (212) 596-9000  
TELEFAX: (212) 596-9090  
TELEX: 14-8367  
INFORMATION FOR SEQ ID NO: 1:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 86 amino acids  
TYPE: amino acid  
TOPOLOGY: linear  
MOLECULE TYPE: protein  
ORIGINAL SOURCE: human immunodeficiency virus  
STRAIN: type 1  
US-08-450-098-1

Query Match 21.6%; Score 486; DB 1; Length 86;  
Best Local Similarity 100.0%; Pred. No. 1.2e-41;  
Matches 85; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
QY 318 EPVDPRLPEPKHPSQPKTACTNCTCYCKKCCFHCQVCFITKALGISYGRKKRRRRPPQG 377  
DB 2 EPVDPRLPEPKHPSQPKTACTNCTCYCKKCCFHCQVCFITKALGISYGRKKRRRRPPQG 61  
QY 378 SOTHQVSLSKOPTSQSRGDPGPK 402  
DB 62 SOTHQVSLSKOPTSQSRGDPGPK 86

RESULT 18  
US-08-451-233-1  
Sequence 1, Application US/08451233  
Patent No. 5747641  
GENERAL INFORMATION:  
APPLICANT: FRANKEL, Alan  
APPLICANT: PABO, Carl  
APPLICANT: BARSOUM, James G.  
APPLICANT: FAWELL, Stephen E.  
APPLICANT: PEPINSKY, R. B.  
TITLE OF INVENTION: TAT-DERIVED TRANSPORT POLYPEPTIDES  
NUMBER OF SEQUENCES: 69  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: FISH & NEAVE  
STREET: 1251 Avenue of the Americas  
CITY: New York  
STATE: New York  
COUNTRY: USA  
ZIP: 10020  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: Patent In Release #1.0, Version #1.25  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/451.233  
FILING DATE: 25-MAY-1995  
CLASSIFICATION: 514  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: US 08/235,403  
FILING DATE: 28-APR-1994  
APPLICATION NUMBER: US 07/934,375  
FILING DATE: 21-AUG-1992  
APPLICATION NUMBER: US 07/098,766  
FILING DATE: 28-JUL-1993  
APPLICATION NUMBER: PCT/US93/07833  
FILING DATE: 19-AUG-1993  
APPLICATION NUMBER: US 07/454,450  
FILING DATE: 21-DEC-1989  
APPLICATION NUMBER: US 07/636,662

```

; FILING DATE: 02-JAN-1991
; APPLICATION NUMBER: US 08/158,015
; FILING DATE: 24-NOV-1993
; ATTORNEY/AGENT INFORMATION:
; NAME: Haley Jr., James F.
; REGISTRATION NUMBER: 27,794
; REFERENCE/DOCKET NUMBER: B170 CIP 2
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (212) 596-9000
; TELEFAX: (212) 596-9090
; TELEX: 14-8367
; INFORMATION FOR SEQ ID NO: 1:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 86 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: protein
; ORIGINAL SOURCE:
; ORGANISM: human immunodeficiency virus
; STRAIN: type 1
; US-08-451-233-1

Query Match          21.6%; Score 486; DB 1; Length 86;
Best Local Similarity 100.0%; Pred. No. 1.2e-41;
Matches 85; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 318 EPVDRLPEPKHPGQPKTACTNCYKCCFHCQVCFITKALGISYGRKKRRRPPG 377
Db 2 EPVDRLPEPKHPGQPKTACTNCYKCCFHCQVCFITKALGISYGRKKRRRPPG 61
QY 378 SOTHQVSLSKQPTSQSRGDPGPK 402
Db 62 SOTHQVSLSKQPTSQSRGDPGPK 86

RESULT 19
US-08-450-236-1
; Sequence 1, Application US/08450236
; Patent No. 5804604
; GENERAL INFORMATION:
; APPLICANT: FRANKEL, Alan
; APPLICANT: PABO, Carl
; APPLICANT: BARSOUM, James G.
; APPLICANT: FAWELL, Stephen E.
; APPLICANT: PEPINSKY, R. B.
; TITLE OF INVENTION: TAT-DERIVED TRANSPORT POLYPEPTIDES
; NUMBER OF SEQUENCES: 69
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: FISH & NEAVE
; STREET: 1251 Avenue of the Americas
; CITY: New York
; STATE: New York
; COUNTRY: USA
; ZIP: 10020
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patent In Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/450,236
; FILING DATE: 25-MAY-1995
; CLASSIFICATION: 514
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/235,403
; FILING DATE: 28-APR-1994
; APPLICATION NUMBER: US 07/934,375
; FILING DATE: 21-AUG-1992
; APPLICATION NUMBER: US 07/098,766
; FILING DATE: 28-JUL-1993
; APPLICATION NUMBER: PCT/US93/07833
; FILING DATE: 19-AUG-1993

; FILING DATE: 02-JAN-1991
; APPLICATION NUMBER: US 07/454,450
; FILING DATE: 21-DEC-1989
; APPLICATION NUMBER: US 07/636,662
; FILING DATE: 02-JAN-1991
; APPLICATION NUMBER: US 08/158,015
; FILING DATE: 24-NOV-1993
; ATTORNEY/AGENT INFORMATION:
; NAME: Haley Jr., James F.
; REGISTRATION NUMBER: 27,794
; REFERENCE/DOCKET NUMBER: B170 CIP 2
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (212) 596-9000
; TELEFAX: (212) 596-9090
; TELEX: 14-8367
; INFORMATION FOR SEQ ID NO: 1:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 86 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: protein
; ORIGINAL SOURCE:
; ORGANISM: human immunodeficiency virus
; STRAIN: type 1
; US-08-450-236-1

Query Match          21.6%; Score 486; DB 1; Length 86;
Best Local Similarity 100.0%; Pred. No. 1.2e-41;
Matches 85; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 318 EPVDRLPEPKHPGQPKTACTNCYKCCFHCQVCFITKALGISYGRKKRRRPPG 377
Db 2 EPVDRLPEPKHPGQPKTACTNCYKCCFHCQVCFITKALGISYGRKKRRRPPG 61
QY 378 SOTHQVSLSKQPTSQSRGDPGPK 402
Db 62 SOTHQVSLSKQPTSQSRGDPGPK 86

RESULT 20
US-07-808-452-1
; Sequence 1, Application US/07808452
; Patent No. 6063612
; GENERAL INFORMATION:
; APPLICANT: Jayasena, Sumedha D.
; APPLICANT: Johnston, Brian H.
; TITLE OF INVENTION: Antiviral Reagents Based on
; TITLE OF INVENTION: RNA-Binding Proteins
; NUMBER OF SEQUENCES: 24
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Law Offices of Peter Dehlinger
; STREET: 350 Cambridge Avenue, Suite 300
; CITY: Palo Alto
; STATE: CA
; COUNTRY: USA
; ZIP: 94306
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patent In Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/07/808,452
; FILING DATE: 19911213
; CLASSIFICATION: 514
; ATTORNEY/AGENT INFORMATION:
; NAME: Fabian, Gary R.
; REGISTRATION NUMBER: 33,875
; REFERENCE/DOCKET NUMBER: 8255-008
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (415) 324-0880
; TELEFAX: (415) 324-0960
; INFORMATION FOR SEQ ID NO: 1:
```

SEQUENCE CHARACTERISTICS:  
LENGTH: 86 amino acids  
TYPE: AMINO ACID  
TOPOLOGY: linear  
MOLECULE TYPE: protein  
HYPOTHETICAL: NO  
ANTI-SENSE: NO  
ORIGINAL SOURCE:  
INDIVIDUAL ISOLATE: the sequence of the TAT protein of  
INDIVIDUAL ISOLATE: HIV-1  
US-07-808-452-1

Query Match 21.6%; Score 486; DB 3; Length 86;  
Best Local Similarity 100.0%; Pred. No. 1.2e-41;  
Matches 85; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 318 EPVDPRLPFWKHPGSPQKPTACTNCYKCCFHCQVCFITKALGISYGRKKRRRRPPOG 377  
|||||

Db 2 EPVDPRLPFWKHPGSPQKPTACTNCYKCCFHCQVCFITKALGISYGRKKRRRRPPOG 61  
|||||

QY 378 SOTHQVLSKQPTSQRGDPGPK 402  
|||||

Db 62 SOTHQVLSKQPTSQRGDPGPK 86  
|||||

## RESULT 21

US-09-030-613-19  
Sequence 19, Application US/09030613  
Patent No. 6083706

## GENERAL INFORMATION:

APPLICANT: FlorKiewicz, Robert Z.

APPLICANT: Baird, J. Andrew

TITLE OF INVENTION: INHIBITORS OF LEADERLESS PROTEIN EXPORT

NUMBER OF SEQUENCES: 36

CORRESPONDENCE ADDRESS:

ADDRESSEE: SEED and BERRY LLP

STREET: 6300 Columbia Center, 701 Fifth Avenue

CITY: Seattle

STATE: Washington

COUNTRY: USA

ZIP: 98104

COMPUTER READABLE FORM:

MEDIUM TYPE: Floppy disk

COMPUTER: IBM PC compatible

OPERATING SYSTEM: PC-DOS/MS-DOS

SOFTWARE: PatentIn Release #1.0, Version #1.30

CURRENT APPLICATION DATA:

APPLICATION NUMBER: US/09/030,613

FILING DATE: 25-FEB-1998

CLASSIFICATION:

ATTORNEY/AGENT INFORMATION:

NAME: No. 6083706tendurg Ph.D., Carol

REGISTRATION NUMBER: 39,317

REFERENCE/DOCKET NUMBER: 760100.418C1

TELECOMMUNICATION INFORMATION:

TELEPHONE: (206) 622-4900

TELEFAX: (206) 682-6031

INFORMATION FOR SEQ ID NO: 19:

SEQUENCE CHARACTERISTICS:

LENGTH: 86 amino acids

TYPE: amino acid

STRANDEDNESS:

TOPOLOGY: linear

US-09-030-613-19

Query Match 21.6%; Score 486; DB 3; Length 86;  
Best Local Similarity 100.0%; Pred. No. 1.2e-41;  
Matches 85; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 318 EPVDPRLPFWKHPGSPQKPTACTNCYKCCFHCQVCFITKALGISYGRKKRRRRPPOG 377  
|||||

Db 2 EPVDPRLPFWKHPGSPQKPTACTNCYKCCFHCQVCFITKALGISYGRKKRRRRPPOG 61  
|||||

QY 378 SOTHQVLSKQPTSQRGDPGPK 402  
|||||

Db 62 SOTHQVLSKQPTSQRGDPGPK 86  
|||||

## RESULT 22

US-09-124-900-6

Sequence 6, Application US/09124900

Patent No. 6268484

GENERAL INFORMATION:

APPLICANT: KATINGER, Hermann

APPLICANT: BUCHACHER, Andrea

APPLICANT: ERNST, Wolfgang

APPLICANT: BALLAUN, Claudia

APPLICANT: PUTSCHER, Martin

APPLICANT: TRKOLA, Alexandra

APPLICANT: PREDL, Renate

APPLICANT: SCHMATZ, Christine

APPLICANT: KLIMA, Annelies

APPLICANT: STEINDL, Franz

APPLICANT: MUSTER, Thomas

TITLE OF INVENTION: HIV-Vaccines

FILE REFERENCE: 1939-112P

CURRENT APPLICATION NUMBER: US/09/124,900

CURRENT FILING DATE: 1998-07-30

PRIOR APPLICATION NUMBER: PCT/EP95/01481

PRIOR FILING DATE: 1995-04-19

NUMBER OF SEQ ID NOS: 11

SOFTWARE: PatentIn version 3.0

SEQ ID NO 6

LENGTH: 86

TYPE: PRT

ORGANISM: Human immunodeficiency virus type 1

US-09-124-900-6

Query Match 21.6%; Score 486; DB 4; Length 86;  
Best Local Similarity 100.0%; Pred. No. 1.2e-41;  
Matches 85; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 318 EPVDPRLPFWKHPGSPQKPTACTNCYKCCFHCQVCFITKALGISYGRKKRRRRPPOG 377  
|||||

Db 2 EPVDPRLPFWKHPGSPQKPTACTNCYKCCFHCQVCFITKALGISYGRKKRRRRPPOG 61  
|||||

QY 378 SOTHQVLSKQPTSQRGDPGPK 402  
|||||

Db 62 SOTHQVLSKQPTSQRGDPGPK 86  
|||||

## RESULT 23

US-09-451-905-19

Sequence 19, Application US/09451905

Patent No. 6306613

GENERAL INFORMATION:

APPLICANT: Robert Z. FlorKiewicz

APPLICANT: Andrew Baird

APPLICANT: Dale E. Warnock

TITLE OF INVENTION: MODULATORS OF LEADERLESS PROTEIN EXPORT

TITLE OF INVENTION: AND METHODS FOR IDENTIFYING AND USING THE SAME

FILE REFERENCE: 200124.402C4

CURRENT APPLICATION NUMBER: US/09/451,905

CURRENT FILING DATE: 1999-12-01

NUMBER OF SEQ ID NOS: 48

SOFTWARE: FastSEQ for Windows Version 4.0

SEQ ID NO 19

LENGTH: 86

TYPE: PRT

ORGANISM: Homo sapien

US-09-451-905-19

Query Match 21.6%; Score 486; DB 4; Length 86;  
Best Local Similarity 100.0%; Pred. No. 1.2e-41;  
Matches 85; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 318 EPVDPRLPEPKHGPSQPKTACTNCKYCKKCFHCQVCFITKALGISYGRKKRRRRPPQG 377  
|||||  
Db 2 EPVDPRLPEPKHGPSQPKTACTNCKYCKKCFHCQVCFITKALGISYGRKKRRRRPPQG 61  
|||||

QY 378 SQTHQVSLSKQPTSQSRGDPGPK 402  
|||||  
Db 62 SQTHQVSLSKQPTSQSRGDPGPK 86  
|||||

RESULT 24  
US-08-235-403-1  
; Sequence 1, Application US/08235403  
; Patent No. 6316003  
; GENERAL INFORMATION:  
; APPLICANT: FRANKEL, Alan  
; APPLICANT: PABO, Carl  
; APPLICANT: BARSOUM, James G.  
; APPLICANT: FAWELL, Stephen E.  
; APPLICANT: PEPINSKY, R. B.  
; TITLE OF INVENTION: TAT-DERIVED TRANSPORT POLYPEPTIDES  
; NUMBER OF SEQUENCES: 69  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: FISH & NEAVE  
; STREET: 1251 Avenue of the Americas  
; CITY: New York  
; STATE: New York  
; COUNTRY: USA  
; ZIP: 10020  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Floppy disk  
; COMPUTER: IBM PC compatible  
; OPERATING SYSTEM: PC-DOS/MS-DOS  
; SOFTWARE: PatentIn Release #1.0, Version #1.25  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/08/235,403  
; FILING DATE:  
; CLASSIFICATION: 514  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: US/934,375  
; FILING DATE: 21-AUG-1992  
; APPLICATION NUMBER: US/098,766  
; FILING DATE: 28-JUL-1993  
; APPLICATION NUMBER: PCT/US93/07833  
; FILING DATE: 19-AUG-1993  
; APPLICATION NUMBER: US/07454,450  
; FILING DATE: 21-DEC-1989  
; APPLICATION NUMBER: US/07636,662  
; FILING DATE: 02-JAN-1991  
; APPLICATION NUMBER: US/08/158,015  
; FILING DATE: 24-NOV-1993  
; ATTORNEY/AGENT INFORMATION:  
; NAME: Haley Jr., James F.  
; REGISTRATION NUMBER: 27,794  
; REFERENCE/DOCKET NUMBER: B170 CIP 2  
; TELEPHONE: (212) 596-9000  
; TELEFAX: (212) 596-9090  
; TELEX: 14-8367  
; INFORMATION FOR SEQ ID NO: 1:  
; SEQUENCE CHARACTERISTICS:  
; LENGTH: 86 amino acids  
; TYPE: amino acid  
; TOPOLOGY: linear  
; MOLECULE TYPE: protein  
; ORIGINAL SOURCE:  
; ORGANISM: human immunodeficiency virus  
; STRAIN: type 1  
US-08-235-403-1

Query Match 21.6%; Score 486; DB 4; Length 86;  
Best Local Similarity 100.0%; Pred. No. 1.2e-41;  
Matches 85; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 318 EPVDPRLPEPKHGPSQPKTACTNCKYCKKCFHCQVCFITKALGISYGRKKRRRRPPQG 377  
|||||  
Db 2 EPVDPRLPEPKHGPSQPKTACTNCKYCKKCFHCQVCFITKALGISYGRKKRRRRPPQG 61  
|||||

QY 378 SQTHQVSLSKQPTSQSRGDPGPK 402  
|||||  
Db 62 SQTHQVSLSKQPTSQSRGDPGPK 86  
|||||

RESULT 25  
PCT-US92-10770-1  
; Sequence 1, Application PC/TUS9210770  
; GENERAL INFORMATION:  
; APPLICANT: Jayasena, Stmedha D.  
; APPLICANT: Johnston, Brian H.  
; TITLE OF INVENTION: Antiviral Reagents Based on  
; TITLE OF INVENTION: RNA-Binding Proteins  
; NUMBER OF SEQUENCES: 24  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: SRI International  
; STREET: 333 Ravenswood Avenue  
; CITY: Menlo Park  
; STATE: CA  
; COUNTRY: USA  
; ZIP: 94025  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Floppy disk  
; COMPUTER: IBM PC compatible  
; OPERATING SYSTEM: PC-DOS/MS-DOS  
; SOFTWARE: PatentIn Release #1.0, Version #1.25  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: PCT/US92/10770  
; FILING DATE: 19921211  
; CLASSIFICATION:  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: US 07/808,452  
; FILING DATE: 13-DEC-1991  
; ATTORNEY/AGENT INFORMATION:  
; NAME: Fabian, Gary R.  
; REGISTRATION NUMBER: 33,875  
; REFERENCE/DOCKET NUMBER: P-2962  
; TELECOMMUNICATION INFORMATION:  
; TELEPHONE: (415) 859-4550  
; TELEFAX: (415) 859-3880  
; INFORMATION FOR SEQ ID NO: 1:  
; SEQUENCE CHARACTERISTICS:  
; LENGTH: 86 amino acids  
; TYPE: AMINO ACID  
; TOPOLOGY: linear  
; MOLECULE TYPE: protein  
; HYPOTHETICAL: NO  
; ANTI-SENSE: NO  
; ORIGINAL SOURCE:  
; INDIVIDUAL ISOLATE: the sequence of the TAT protein of  
; INDIVIDUAL ISOLATE: HIV-1  
PCT-US92-10770-1

Query Match 21.6%; Score 486; DB 5; Length 86;  
Best Local Similarity 100.0%; Pred. No. 1.2e-41;  
Matches 85; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 318 EPVDPRLPEPKHGPSQPKTACTNCKYCKKCFHCQVCFITKALGISYGRKKRRRRPPQG 377  
|||||  
Db 2 EPVDPRLPEPKHGPSQPKTACTNCKYCKKCFHCQVCFITKALGISYGRKKRRRRPPQG 61  
|||||

QY 378 SQTHQVSLSKQPTSQSRGDPGPK 402  
|||||



Db 62 SOTHQVSLSKQPTSQSRGDPGPK 86

RESULT 26  
PCT-US95-06077-2  
Sequence 2, Application PC/TUS9506077  
GENERAL INFORMATION:  
APPLICANT: Immunobiology Research, Institute Inc.  
TITLE OF INVENTION: Vaccine Interdiction of Extracellular  
TITLE OF INVENTION: Transactivating Proteins of Human Immunodeficiency Virus  
TITLE OF INVENTION: and Other Chronically Infecting Viruses Employing Similar  
TITLE OF INVENTION: Intercellular Transactivating Strategies  
NUMBER OF SEQUENCES: 38  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: Howson and Howson  
STREET: Spring House Corporate Cntr, PO Box 457  
CITY: Spring House  
STATE: Pennsylvania  
COUNTRY: USA  
ZIP: 19477  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: PatentIn Release #1.0, Version #1.25  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: PCT/US95/06077  
FILING DATE:  
CLASSIFICATION:  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: US 08/247,991  
FILING DATE: 23-MAY-1994  
ATTORNEY/AGENT INFORMATION:  
NAME: Bak, Mary E.  
REGISTRATION NUMBER: 31,215  
REFERENCE/DOCKET NUMBER: IRI44PCT  
TELEPHONE: 215-540-9200  
TELEFAX: 215-540-5818  
INFORMATION FOR SEQ ID NO: 2:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 86 amino acids  
TYPE: amino acid  
TOPOLOGY: linear  
MOLECULE TYPE: protein  
PCT-US95-06077-2

Query Match 21.6%; Score 486; DB 5; Length 86;  
Best Local Similarity 100.0%; Pred. No. 1.2e-41;  
Matches 85; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 318 EPVDPLEPKHPSQPKTACTNCKKCFHCQVCFITKALGISYGRKKRRRPPQG 377  
|||||  
Db 2 EPVDPLEPKHPSQPKTACTNCKKCFHCQVCFITKALGISYGRKKRRRPPQG 61  
|||||

QY 378 SOTHQVSLSKQPTSQSRGDPGPK 402  
|||||  
Db 62 SOTHQVSLSKQPTSQSRGDPGPK 86

RESULT 27  
US-08-815-809-5  
Sequence 5, Application US/08815809  
Patent No. 6004777  
GENERAL INFORMATION:  
APPLICANT: TARTAGLIA, James  
APPLICANT: GOEBEL, Scott J.  
APPLICANT: COX, William I.  
APPLICANT: GETTIG, Russell R.  
APPLICANT: PINCUS, Steven E.  
APPLICANT: PAOLETTI, Enzo  
APPLICANT: JACOBS, Bertram L.

; TITLE OF INVENTION: VECTORS HAVING ENHANCED EXPRESSION, AND METHODS OF  
; TITLE OF INVENTION: MAKING AND USES THEREOF  
; FILE REFERENCE: 454310-3010  
; CURRENT APPLICATION NUMBER: US/08/815.809  
; CURRENT FILING DATE: 1997-03-12  
; NUMBER OF SEQ ID NOS: 23  
; SOFTWARE: PatentIn Ver. 2.0  
; SEQ ID NO 5  
; LENGTH: 266  
; TYPE: PRT  
; ORGANISM: Vaccinia virus  
US-08-815-809-5

Query Match 21.2%; Score 478.5; DB 3; Length 266;  
Best Local Similarity 77.3%; Pred. No. 3.5e-40;  
Matches 92; Conservative 4; Mismatches 6; Indels 17; Gaps 4;

QY 175 VGFPTVPOVPLRPMYKAAVDLSHFLKKEGGLTHSRRQDILDLYHTQGYPPDQ 234  
|||||  
Db 136 VGFPTVPOVPLRPMYKAAVDLSHFLKKEGGLTHSRRQDILDLYHTQGYPPDQ 195  
|||||

QY 235 NYTPGPGVRYPLTFGWCYKLVPEPKVEANKGENTSLHVPVSLH-GMDDPEREYLEW 292  
|||||  
Db 196 NYTPGPGVRYPLTFGWCYKLVPM----IETV-----PVKLPKGMGDP--KVQW 238  
|||||

RESULT 28  
US-08-816-155B-43  
Sequence 43, Application US/08816155B  
Patent No. 5990091  
GENERAL INFORMATION:  
APPLICANT: TARTAGLIA, JAMES  
APPLICANT: COX, WILLIAM I.  
APPLICANT: GETTIG, RUSSELL R.  
APPLICANT: MARTINEZ, HECTOR  
APPLICANT: PAOLETTI, ENZO  
APPLICANT: PINCUS, STEVEN E.  
TITLE OF INVENTION: VECTORS HAVING ENHANCED EXPRESSION, AND  
TITLE OF INVENTION: METHODS OF MAKING AND USES THEREOF  
NUMBER OF SEQUENCES: 48  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: FROMMER LAWRENCE & HAUG LLP  
STREET: 745 FIFTH AVENUE  
CITY: NEW YORK  
STATE: NEW YORK  
COUNTRY: USA  
ZIP: 10151  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: PatentIn Release #1.0, Version #1.30  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/816.155B  
FILING DATE: 12-MAR-1997  
CLASSIFICATION: 514  
ATTORNEY/AGENT INFORMATION:  
NAME: KOWALSKI, THOMAS J.  
REGISTRATION NUMBER: 32,147  
REFERENCE/DOCKET NUMBER: 454310-2990  
TELEPHONE: 212-588-0800  
TELEFAX: 212-588-0500  
INFORMATION FOR SEQ ID NO: 43:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 280 amino acids  
TYPE: amino acid  
STRANDEDNESS: n/a  
TOPOLOGY: linear  
MOLECULE TYPE: amino acid  
US-08-816-155B-43

Query Match 21.2%; Score 478.5; DB 2; Length 280;  
Best Local Similarity 77.3%; Pred. No. 3.7e-40;  
Matches 92; Conservative 4; Mismatches 6; Indels 17; Gaps 4;

QY 175 VGFPVTPQVPLRPMTYKAAVDLSHFLKEKGLGGLHSQRRQDILDLYHTQGYFPDQW 234  
|||||  
DB 150 VGFPVTPQVPLRPMTYKAAVDLSHFLKEKGLGGLHSQRRQDILDLYHTQGYFPDQW 209  
|||||

QY 235 NYTPGPGVRYPLTFGWCYKLVPEPDKVEANKGENTSLLLHPVSLH-GMDDPEREVLW 292  
|||||  
DB 210 NYTPGPGVRYPLTFGWCYKLVPEPDKVEANKGENTSLLLHPVSLH-GMDDPEREVLW 252  
|||||

RESULT 29  
US-09-079-587-43  
; Sequence 43, Application US/09079587  
; Patent No. 6130066  
; GENERAL INFORMATION:  
; APPLICANT: TARTAGLIA, JAMES  
; APPLICANT: COX, WILLIAM I.  
; APPLICANT: GETTIG, RUSSELL R.  
; APPLICANT: MARTINEZ, HECTOR  
; APPLICANT: PAOLETTI, ENZO  
; APPLICANT: PINCUS, STEVEN E.  
; TITLE OF INVENTION: VECTORS HAVING ENHANCED EXPRESSION, AND  
; TITLE OF INVENTION: METHODS OF MAKING AND USES THEREOF  
; NUMBER OF SEQUENCES: 48  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: FROMMER LAWRENCE & HAUG LLP  
; STREET: 745 FIFTH AVENUE  
; CITY: NEW YORK  
; STATE: NEW YORK  
; COUNTRY: USA  
; ZIP: 10151  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Floppy disk  
; COMPUTER: IBM PC compatible  
; OPERATING SYSTEM: PC-DOS/MS-DOS  
; SOFTWARE: PatentIn Release #1.0, Version #1.30  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/09/079,587  
; FILING DATE:  
; CLASSIFICATION:  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: US 08/816,155  
; FILING DATE: 12-MAR-1997  
; ATTORNEY/AGENT INFORMATION:  
; NAME: KOHALSKI, THOMAS J.  
; REGISTRATION NUMBER: 32,147  
; REFERENCE/DOCKET NUMBER: 454310-2990  
; TELECOMMUNICATION INFORMATION:  
; TELEPHONE: 212-588-0800  
; TELEFAX: 212-588-0500  
; INFORMATION FOR SEQ ID NO: 43:  
; SEQUENCE CHARACTERISTICS:  
; LENGTH: 280 amino acids  
; TYPE: amino acid  
; STRANDEDNESS: n/a  
; TOPOLOGY: linear  
; MOLECULE TYPE: amino acid  
US-09-079-587-43

Query Match 21.2%; Score 478.5; DB 4; Length 280;  
Best Local Similarity 77.3%; Pred. No. 3.7e-40;  
Matches 92; Conservative 4; Mismatches 6; Indels 17; Gaps 4;

QY 175 VGFPVTPQVPLRPMTYKAAVDLSHFLKEKGLGGLHSQRRQDILDLYHTQGYFPDQW 234  
|||||  
DB 150 VGFPVTPQVPLRPMTYKAAVDLSHFLKEKGLGGLHSQRRQDILDLYHTQGYFPDQW 209  
|||||

QY 235 NYTPGPGVRYPLTFGWCYKLVPEPDKVEANKGENTSLLLHPVSLH-GMDDPEREVLW 292  
|||||

DB 210 NYTPGPGVRYPLTFGWCYKLVPEPDKVEANKGENTSLLLHPVSLH-GMDDPEREVLW 252  
|||||

RESULT 30  
US-08-505-210-1  
; Sequence 1, Application US/08505210  
; Patent No. 5981258  
; GENERAL INFORMATION:  
; APPLICANT: MEHTALI, Majid  
; APPLICANT: GUSS, Tania  
; TITLE OF INVENTION: COMPOSITION OF TRANS-DOMINANT VARIANTS  
; TITLE OF INVENTION: OF VIRAL PROTEINS FOR OBTAINING AN ANTIVIRAL EFFECT  
; NUMBER OF SEQUENCES: 7  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: BURNS, DOANE, SWECKER & MATHIS, L.L.P.  
; STREET: 1737 King Street, Suite 500  
; CITY: Alexandria  
; STATE: Virginia  
; COUNTRY: United States  
; ZIP: 22314-2756  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Floppy disk  
; COMPUTER: IBM PC compatible  
; OPERATING SYSTEM: PC-DOS/MS-DOS  
; SOFTWARE: PatentIn Release #1.0, Version #1.30  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/08/505,210  
; FILING DATE: 14-AUG-1995  
; CLASSIFICATION: 435  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: US 08/215,248  
; FILING DATE: 21-MAR-1994  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: WO PCT/FR94/01457  
; FILING DATE: 13-DEC-1994  
; ATTORNEY/AGENT INFORMATION:  
; NAME: Dadio, Susan M.  
; REGISTRATION NUMBER: 40,373  
; REFERENCE/DOCKET NUMBER: 017753-066  
; TELECOMMUNICATION INFORMATION:  
; TELEPHONE: (703) 836-6620  
; TELEFAX: (703) 836-2021  
; INFORMATION FOR SEQ ID NO: 1:  
; SEQUENCE CHARACTERISTICS:  
; LENGTH: 86 amino acids  
; TYPE: amino acid  
; STRANDEDNESS: single  
; TOPOLOGY: linear  
; MOLECULE TYPE: protein  
; HYPOTHETICAL: NO  
; ANTI-SENSE: NO  
; ORIGINAL SOURCE:  
; ORGANISM: Human immunodeficiency virus type 1  
; STRAIN: Lai  
; INDIVIDUAL ISOLATE: sequence of the TAT protein of HIV-1  
US-08-505-210-1

Query Match 20.8%; Score 470; DB 2; Length 86;  
Best Local Similarity 96.5%; Pred. No. 4.8e-40;  
Matches 82; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

QY 318 EPVDPRLPEPKHPSQPKTACTNCKYCKCFHCOVCFITKALGTSYGRKKRRRRPPQG 377  
|||||

DB 2 EPVDPRLPEPKHPSQPKTACTNCKYCKCFHCOVCFITKALGTSYGRKKRRRRPPQG 61  
|||||

QY 378 SQTHQVSLSKOPTSQSRGDPDTPGKE 402  
|||||

DB 62 SQTHQVSLSKOPTSQSRGDPDTPGKE 86  
|||||

Mon Aug 26 09:05:50 2002

us-09-509-239-17.rai

Page 14

Search completed: August 26, 2002, 08:10:40  
Job time: 129 sec

GenCore version 4.5  
Copyright (c) 1993 - 2000 Compugen Ltd.

OM protein - protein search, using sw model

Run on: August 26, 2002, 08:13:21 ; Search time 50.45 Seconds  
(without alignments)  
782.809 Million cell updates/sec

Title: US-09-509-239-17  
Perfect score: 2255  
Sequence: 1 CSSHSSNANTQMSKDKIII.....QSRGDPGPKETSGHHHHH 411

Scoring table: BLOSUM62  
Gapop 10.0 , Gapext 0.5.

Searched: 283138 seqs, 96089334 residues

Total number of hits satisfying chosen parameters: 283138

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : PIR71:\*\*

1: pir1:\*\*

2: pir2:\*\*

3: pir3:\*\*

4: pir4:\*\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

#### SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	1130	50.1	206	1 ASLJFV	nef protein - huma
2	1115	49.4	205	2 S43467	nef protein - huma
3	1115	49.4	206	2 S03244	nef protein (clone
4	1114	49.4	206	1 ASLJVL	nef protein - huma
5	1108	49.1	206	1 ASLJ12	nef protein - huma
6	1097	48.6	206	2 JC3400	nef protein (clone
7	1096	48.6	206	2 S03245	nef protein - huma
8	1096	48.6	206	2 S33986	nef protein - huma
9	1081.5	48.0	205	1 ASLJH3	nef protein - huma
10	1019	45.2	204	2 S24985	nef protein - huma
11	1019	45.2	218	1 ASLJBR	nef protein - huma
12	1007	44.7	206	2 S25937	nef protein - huma
13	997	44.2	210	1 I44J02	nef protein - huma
14	991.5	44.0	214	1 I44J01	nef protein - huma
15	987	43.8	182	2 S03247	nef protein (clone
16	987	43.8	204	2 S03246	nef protein (clone
17	911	40.4	212	1 QQLJZR	nef protein - huma
18	903.5	40.1	207	1 QQLJND	nef protein - huma
19	837	37.1	205	1 B44963	nef protein - huma
20	834.5	37.0	209	2 T01673	nef protein - huma
21	832	36.9	205	1 ASLJIK	nef protein - simi
22	578.5	25.7	364	2 S59931	glycerophosphodies
23	578.5	25.7	364	2 S59932	glycerophosphodies
24	578.5	25.7	364	2 A43576	glycerophosphodies
25	578.5	25.7	364	2 S59934	glycerophosphodies
26	572.5	25.4	364	2 S59933	glycerophosphodies
27	562.5	24.9	364	2 G64086	glycerophosphodies
28	556.5	24.7	364	2 S59936	glycerophosphodies
29	486	21.6	86	2 S33982	trans-activating t

30 478 21.2 95 1 TNLJ12 trans-activating t  
31 470 20.8 86 2 A25700 trans-activating t  
32 443 19.6 226 2 S46353 nef protein - simi  
33 439.5 19.5 97 2 S54385 nef protein - huma  
34 437 19.4 101 2 T09446 tat protein - huma  
35 434 19.2 101 1 E44001 trans-activating t  
36 429.5 19.0 309 2 S07993 nef protein - simi  
37 426.5 18.9 259 2 T11567 nef protein - simi  
38 421.5 18.7 255 2 S53099 nef protein - huma  
39 421 18.7 257 2 S12160 nef protein - huma  
40 413 18.3 260 1 ASLJJCZ nef protein - huma  
41 411.5 18.2 263 2 T11556 nef protein - simi  
42 411 18.2 86 2 JC5591 transactivator pro  
43 409 18.1 250 2 S54851 nef protein - simi  
44 408 18.1 86 2 S54381 tat protein - huma  
45 405 18.0 86 1 TNLJZR trans-activating t

#### ALIGNMENTS

RESULT 1

ASLJFV

nef protein - human immunodeficiency virus type 1 (isolate LAV-1a)

N;Alternate names: 3'-orf protein; orf-F protein

C:Species: human immunodeficiency virus type 1, HIV-1

A:Note: Host Homo sapiens (man)

C:Date: 17-May-1985 #sequence\_revision 17-May-1985 #text\_change 16-Jul-1999

C:Accession: A04008; S14609

R:Wain-Hobson, S.; Sonigo, P.; Danos, O.; Cole, S.; Allison, M.

Cell 40, 9-17, 1985

A:Title: Nucleotide sequence of the AIDS virus, LAV.

A:Reference number: A90866; MUID:85099333

A:Accession: A04008

A:Molecule type: DNA

A:Residues: 1-206 <NA>

A:Cross-references: GB:K02013; NID:g326417; PIDN:AAB59752.1; PID:g326425

A:Experimental source: isolate LAV-1a

R:Ciccarelli, R.B.

Submitted to the EMBL Data Library, March 1991

A:Reference number: S14607

A:Accession: S14609

A>Status: preliminary

A:Molecule type: DNA

A:Residues: 1-206 <CIC>

A:Cross-references: EMBL:X58780; NID:g60113; PIDN:CAA41585.1; PID:g60114

C:Genetics:

A:Gene: nef; 3'-orf; orf-F

C:Superfamily: AIDS nef protein

C:Keywords: AIDS; Immunodeficiency; phosphoprotein

Query Match

Best Local Similarity 50.1%; Score 1130; DB 1; Length 206;

Matches 206; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 110 MGKWSKSSVVGWPTVRMRRAEPAADGVGAASRDLEKKGAITSSNTAATNAACAWLEA 169

Db 1 MGKWSKSSVVGWPTVRMRRAEPAADGVGAASRDLEKKGAITSSNTAATNAACAWLEA 60

QY 170 QEEEEVGFPVTPQVLRPMYTKAADVLSHFLKEKGGLEGLIHSORRODILDLYHTFGY 229

Db 61 QEEEEVGFPVTPQVLRPMYTKAADVLSHFLKEKGGLEGLIHSORRODILDLYHTFGY 120

QY 230 FPDQNTYTPGVRYPITFGWCYKLYPVEPKVEEANKGENTSLHHPVSLHGMDPPREV 289

Db 121 FPDQNTYTPGVRYPITFGWCYKLYPVEPKVEEANKGENTSLHHPVSLHGMDPPREV 180

QY 290 LEWRFSRLAFHFAHVELHPEYFKNC 315

Db 181 LEWRFSRLAFHFAHVELHPEYFKNC 206

```
RESULT 2
S43467
nef protein - human immunodeficiency virus type 1
C:Species: human immunodeficiency virus type 1, HIV-1
C:Date: 20-Oct-1994 #sequence_revision 10-Nov-1995 #text_change 03-May-1996
C:Accession: S43467
R:Freund, J.; Kellner, R.; Houthaeve, T.; Kalbitzer, H.R.
Eur. J. Biochem. 221, 811-819, 1994
A:Title: Stability and proteolytic domains of Nef protein from human immunodeficiency vi
A:Reference: S43467; MUID:94229079
A:Accession: S43467
A:Status: preliminary
A:Molecule type: protein
A:Residues: 1-205 <PRE>
C:Superfamily: AIDS nef protein

Query Match 49.4%; Score 1115; DB 2; Length 205;
Best Local Similarity 99.0%; Pred. No. 1.7e-79;
Matches 203; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 111 GKGWSSVVGWPTVRRMRRAEPAADGVGAASRDLEKHGATSSNTAATNAACAWLEA 170
|||||
Db 1 GKGWSSVVGWPAVRMRRAEPAADGVGAASRDLEKHGATSSNTAATNAACAWLEA 60
|||||

QY 171 EEEVGFPPVTPQVPLRPMTYKAAVDLSHFLKKGGLHSGRRQDILDLWIYHTQGYF 230
|||||
Db 61 EEEVGFPPVTPQVPLRPMTYKAAVDLSHFLKKGGLHSGRRQDILDLWIYHTQGYF 120
|||||

QY 231 PDQNTPTGPGVRYPLTFGCWYKLVPEPDKVEEANKGENTSLHPVSLHGMDPDEREV 290
|||||
Db 121 PDQNTPTGPGVRYPLTFGCWYKLVPEPDKVEEANKGENTSLHPVSLHGMDPDEREV 180
|||||

QY 291 EWRFDLSRLAFHHVARELHPEYFKNC 315
|||||
Db 181 EWRFDLSRLAFHHVARELHPEYFKNC 205
|||||

RESULT 3
S03244
nef protein (clone HXB2) - human immunodeficiency virus type 1
N:Alternate names: 3'-orf protein
C:Species: human immunodeficiency virus type 1, HIV-1
C:Date: 28-Feb-1990 #sequence_revision 28-Feb-1990 #text_change 30-Sep-1993
C:Accession: S03244
R:Ratner, L.; Starcich, B.; Josephs, S.F.; Hahn, B.H.; Reddy, E.P.; Livak, K.J.; Pettawa
Nucleic Acids Res. 13, 8219-8229, 1985
A:Title: Polymorphism of the 3' open reading frame of the virus associated with the acqu
A:Reference number: S03244; MUID:86067228
A:Accession: S03244
A:Molecule type: DNA
A:Residues: 1-206 <RAT>
A:Cross-references: EMBL:X03187
A:Note: the in-frame stop codon at residue 124, shown as 'X', may be suppressed
C:Genetics:
A:Gene: nef; 3'-orf; orf-F
C:Superfamily: AIDS nef protein

Query Match 49.4%; Score 1115; DB 2; Length 206;
Best Local Similarity 98.5%; Pred. No. 1.7e-79;
Matches 203; Conservative 2; Mismatches 1; Indels 0; Gaps 0;

QY 110 MGKWSKSSVVGWPTVRRMRRAEPAADGVGAASRDLEKHGATSSNTAATNAACAWLEA 169
|||||
Db 1 MGKWSKSSVVGWPTVRRMRRAEPAADGVGAASRDLEKHGATSSNTAATNAACAWLEA 60
|||||

QY 170 QEEVGFPPVTPQVPLRPMTYKAAVDLSHFLKKGGLHSGRRQDILDLWIYHTQGY 229
|||||
Db 61 QEEVGFPPVTPQVPLRPMTYKAAVDLSHFLKKGGLHSGRRQDILDLWIYHTQGY 120
|||||

QY 230 FPDQNTPTGPGVRYPLTFGCWYKLVPEPDKVEEANKGENTSLHPVSLHGMDPDEREV 289
|||||
Db 121 FPDQNTPTGPGVRYPLTFGCWYKLVPEPDKVEEANKGENTSLHPVSLHGMDPDEREV 180
|||||

RESULT 4
ASLJVL
nef protein - human immunodeficiency virus type 1 (isolate LV)
N:Alternate names: 3'-orf protein; orf-F protein
C:Species: human immunodeficiency virus type 1, HIV-1
C:Date: 17-May-1985 #sequence_revision 17-May-1985 #text_change 16-Jul-1999
C:Accession: A04007
R:Muesing, M.A.; Smith, D.H.; Cabradilla, C.D.; Benton, C.V.; Lasky, L.A.; Capon, D.
Nature 313, 450-458, 1985
A:Title: Nucleic acid structure and expression of the human AIDS/lymphadenopathy ret
A:Reference number: A93355; MUID:85111157
A:Accession: A04007
A:Molecule type: DNA
A:Residues: 1-206 <MUE>
A:Cross-references: GB:K02083; NID:9555008; PIDN:AAB59874.1; PID:9328560
C:Genetics:
A:Gene: nef; 3'-orf; orf-F
C:Superfamily: AIDS nef protein
C:Keywords: AIDS; immunodeficiency

Query Match 49.4%; Score 1114; DB 1; Length 206;
Best Local Similarity 97.6%; Pred. No. 2.1e-79;
Matches 201; Conservative 3; Mismatches 2; Indels 0; Gaps 0;

QY 110 MGKWSKSSVVGWPTVRRMRRAEPAADGVGAASRDLEKHGATSSNTAATNAACAWLEA 169
|||||
Db 1 MGKWSKSSVVGWPAVRMRRAEPAADGVGAASRDLEKHGATSSNTAATNAACAWLEA 60
|||||

QY 170 QEEVGFPPVTPQVPLRPMTYKAAVDLSHFLKKGGLHSGRRQDILDLWIYHTQGY 229
|||||
Db 61 QEEVGFPPVTPQVPLRPMTYKAAVDLSHFLKKGGLHSGRRQDILDLWIYHTQGY 120
|||||

QY 230 FPDQNTPTGPGVRYPLTFGCWYKLVPEPDKVEEANKGENTSLHPVSLHGMDPDEREV 289
|||||
Db 121 FPDQNTPTGPGVRYPLTFGCWYKLVPEPDKVEEANKGENTSLHPVSLHGMDPDEREV 180
|||||

QY 290 LEWRFDLSRLAFHHVARELHPEYFKNC 315
|||||
Db 181 LEWRFDLSRLAFHHVARELHPEYFKNC 206
|||||

RESULT 5
ASLJ12
nef protein - human immunodeficiency virus type 1 (isolate HTLV-III, 12)
N:Alternate names: 3'-orf protein; orf-F protein
C:Species: human immunodeficiency virus type 1, HIV-1
A:Note: host Homo sapiens (man)
C:Date: 04-Dec-1986 #sequence_revision 31-Dec-1991 #text_change 16-Jul-1999
C:Accession: A04006
R:Arya, S.K.; Gallo, R.C.
Proc. Natl. Acad. Sci. U.S.A. 83, 2209-2213, 1986
A:Title: Three novel genes of human T-lymphotropic virus type III: immune reactivit
A:Reference number: A94093; MUID:86177573
A:Accession: A04006
A:Molecule type: DNA
A:Residues: 1-206 <ARY>
A:Cross-references: EMBL:M11840; NID:g328453; PIDN:AAA45001.1; PID:g328458
C:Genetics:
A:Gene: nef; 3'-orf; orf-F
C:Superfamily: AIDS nef protein
C:Keywords: AIDS; immunodeficiency

Query Match 49.1%; Score 1108; DB 1; Length 206;
Best Local Similarity 97.1%; Pred. No. 6e-79;
```

Matches 200; Conservative 4; Mismatches 2; Indels 0; Gaps 0;

QY 110 MGGKWSKSSVVGWPTVRRMRRAEPAADGVGAASRDLEKKGAISSNTAATNAACAWLEA 169  
 Db 1 MGGKWSKSSVVGWPAVRMRRAEPAADGVGAASRDLEKKGAISSNTAATNAACAWLEA 60

QY 170 QEEVEGFPVTPQVPLRPMTYKAAVDLSHFLKEKGGLEGLIHSQRRQDILDLIWYHTQGY 229  
 Db 61 QEEVEGFPVTPQVPLRPMTYKAAVDLSHFLKEKGGLEGLIHSQRRQDILDLIWYHTQGY 120

QY 230 FPDQNTYTPGPGVRYPLTFGWCYKLVPEPKVEEANKGENTSLHPVSLHGMDDPEREV 289  
 Db 121 FPDQNTYTPGPGVRYPLTFGWCYKLVPEPKVEEANKGENTSLHPVSLHGMDDPEREV 180

QY 290 LEWRDLSRLAFHHVARELHPEYFKNC 315  
 Db 181 LEWRDLSRLAFHHVARELHPEYFKNC 206

RESULT 6  
 JC5400  
 nef protein - Human immunodeficiency virus type 1, HIV-1  
 C:Species: Human immunodeficiency virus type 1, HIV-1  
 C:Date: 07-Jul-1997 #sequence\_revision 18-Jul-1997 #text\_change 31-Oct-1997  
 C:Accession: JC5400  
 R:Macreadie, I.G.; Lowe, M.G.; Curtin, C.C.; Hewish, D.; Azad, A.A.  
 Biochem. Biophys. Res. Commun. 232, 707-711, 1997  
 A:Title: Cytotoxicity resulting from addition of HIV-1 Nef N-terminal peptides to yeast  
 A:Reference number: JC5400; MUID:97271389  
 A:Molecule type: protein  
 A:Residues: 1-206 <MAC>  
 C:Comment: The amino-terminal part possesses membrane-perturbing and fusogenic activities.  
 C:Superfamily: AIDS nef protein  
 F:2-206/Product: nef protein #status predicted <MAT>

Query Match 48.6%; Score 1097; DB 2; Length 206;  
 Best Local Similarity 97.1%; Pred. No. 4.3e-78;  
 Matches 200; Conservative 1; Mismatches 5; Indels 0; Gaps 0;

QY 110 MGGKWSKSSVVGWPTVRRMRRAEPAADGVGAASRDLEKKGAISSNTAATNAACAWLEA 169  
 Db 1 MGGKWSKSSVVGWPAVRMRRAEPAADGVGAASRDLEKKGAISSNTAATNAACAWLEA 60

QY 170 QEEVEGFPVTPQVPLRPMTYKAAVDLSHFLKEKGGLEGLIHSQRRQDILDLIWYHTQGY 229  
 Db 61 QEEVEGFPVTPQVPLRPMTYKAAVDLSHFLKEKGGLEGLIHSQRRQDILDLIWYHTQGY 120

QY 230 FPDQNTYTPGPGVRYPLTFGWCYKLVPEPKVEEANKGENTSLHPVSLHGMDDPEREV 289  
 Db 121 FPDQNTYTPGPGVRYPLTFGWCYKLVPEPKVEEANKGENTSLHPVSLHGMDDPEREV 180

QY 290 LEWRDLSRLAFHHVARELHPEYFKNC 315  
 Db 181 LEWRDLSRLAFHHVARELHPEYFKNC 206

RESULT 7  
 S03245  
 nef protein (clone HXB3) - human immunodeficiency virus type 1  
 N:Alternate names: 3'-orf protein  
 C:Species: human immunodeficiency virus type 1, HIV-1  
 C:Date: 28-Feb-1990 #sequence\_revision 28-Feb-1990 #text\_change 20-Sep-1999  
 C:Accession: S03245  
 R:Ratner, L.; Starcich, B.; Josephs, S.F.; Hahn, B.H.; Reddy, E.P.; Livak, K.J.; Pettawa  
 Nucleic Acids Res. 13, 8219-8229, 1985  
 A:Title: Polymorphism of the 3' open reading frame of the virus associated with the acqu  
 A:Reference number: S03244; MUID:86067228  
 A:Accession: S03245  
 A:Molecule type: DNA  
 A:Residues: 1-206 <RAT>

A:Cross-references: EMBL:X03188; NID:g61556; PIDN:CAA26947.1; PID:g61557  
 C:Genetics:  
 A:Gene: nef; 3'-orf; Orf-F  
 C:Superfamily: AIDS nef protein

Query Match 48.6%; Score 1096; DB 2; Length 206;  
 Best Local Similarity 96.6%; Pred. No. 5.2e-78;  
 Matches 199; Conservative 4; Mismatches 3; Indels 0; Gaps 0;

QY 110 MGGKWSKSSVVGWPTVRRMRRAEPAADGVGAASRDLEKKGAISSNTAATNAACAWLEA 169  
 Db 1 MGGKWSKSSVVGWPAVRMRRAEPAADGVGAASRDLEKKGAISSNTAATNAACAWLEA 60

QY 170 QEEVEGFPVTPQVPLRPMTYKAAVDLSHFLKEKGGLEGLIHSQRRQDILDLIWYHTQGY 229  
 Db 61 QEEVEGFPVTPQVPLRPMTYKAAVDLSHFLKEKGGLEGLIHSQRRQDILDLIWYHTQGY 120

QY 230 FPDQNTYTPGPGVRYPLTFGWCYKLVPEPKVEEANKGENTSLHPVSLHGMDDPEREV 289  
 Db 121 FPDQNTYTPGPGVRYPLTFGWCYKLVPEPKVEEANKGENTSLHPVSLHGMDDPEREV 180

QY 290 LEWRDLSRLAFHHVARELHPEYFKNC 315  
 Db 181 LEWRDLSRLAFHHVARELHPEYFKNC 206

RESULT 8  
 S33986  
 nef protein - human immunodeficiency virus type 1  
 C:Species: human immunodeficiency virus type 1, HIV-1  
 C:Date: 19-Mar-1997 #sequence\_revision 19-Mar-1997 #text\_change 20-Sep-1999  
 C:Accession: S33986  
 R:Carlini, F.  
 submitted to the EMBL Data Library, November 1991  
 A:Reference number: S33979  
 A:Accession: S33986  
 A:Status: preliminary  
 A:Molecule type: mRNA  
 A:Residues: 1-206 <CAR>  
 A:Cross-references: EMBL:Z11530; NID:g60192; PIDN:CAA77629.1; PID:g60200  
 C:Superfamily: AIDS nef protein

Query Match 48.6%; Score 1096; DB 2; Length 206;  
 Best Local Similarity 96.6%; Pred. No. 5.2e-78;  
 Matches 199; Conservative 3; Mismatches 4; Indels 0; Gaps 0;

QY 110 MGGKWSKSSVVGWPTVRRMRRAEPAADGVGAASRDLEKKGAISSNTAATNAACAWLEA 169  
 Db 1 MGGKWSKSSVVGWPAVRMRRAEPAADGVGAASRDLEKKGAISSNTAATNAACAWLEA 60

QY 170 QEEVEGFPVTPQVPLRPMTYKAAVDLSHFLKEKGGLEGLIHSQRRQDILDLIWYHTQGY 229  
 Db 61 QEEVEGFPVTPQVPLRPMTYKAAVDLSHFLKEKGGLEGLIHSQRRQDILDLIWYHTQGY 120

QY 230 FPDQNTYTPGPGVRYPLTFGWCYKLVPEPKVEEANKGENTSLHPVSLHGMDDPEREV 289  
 Db 121 FPDQNTYTPGPGVRYPLTFGWCYKLVPEPKVEEANKGENTSLHPVSLHGMDDPEREV 180

QY 290 LEWRDLSRLAFHHVARELHPEYFKNC 315  
 Db 181 LEWRDLSRLAFHHVARELHPEYFKNC 206

RESULT 9  
 ASLJH3  
 nef protein - human immunodeficiency virus type 1 (isolate HTLV-III, BH10)  
 N:Alternate names: 3'-orf protein; Orf-F protein  
 C:Species: human immunodeficiency virus type 1, HIV-1  
 A:Note: host Homo sapiens (man)  
 C:Date: 17-May-1985 #sequence\_revision 17-May-1985 #text\_change 31-Jan-1997  
 C:Accession: A04005

R; Ratner, L.; Haseltine, W.; Patarca, R.; Livak, K.J.; Starcich, B.; Josephs, S.F.; Doria  
nberger, J.A.; Papas, T.S.; Graybe, J.; Chang, N.F.; Gallo, R.C.; Wong-Staal, F.

Nature 313, 277-284, 1985

A: Title: Complete nucleotide sequence of the AIDS virus, HTLV-III.

A: Reference number: A93353; MUID: 85111123

A: Accession: A04005

A: Molecule type: DNA

A: Residues: 1-205 <RAT>

C: Genetics:

A: Gene: nef; 3'-orf; orf-F

C: Superfamily: AIDS nef protein

C: Keywords: AIDS; immunodeficiency

Query Match 48.0%; Score 1081.5; DB 1; Length 205;  
Best Local Similarity 96.1%; Pred. No. 6.9e-77;  
Matches 198; Conservative 4; Mismatches 3; Indels 1; Gaps 1;

QY 110 MGGKWSKSVVGVPTVPRMRRAEPAADGVGAASRDLEKKGALITSSNTAATNAACAWLEA 169

Db 1 MGGKWSKSVVGVPTVPRMRRAEPAADGVGAASRDLEKKGALITSSNTAATNAACAWLEA 60

QY 170 QEEVEGFPVTPQVPLRPMTYKAAVDLSHFLKEKGLGLIHSQRQDILDWIYHTQGY 229

Db 61 QEEVEGFPVTPQVPLRPMTYKAAVDLSHFLKEKGLGLIHSQRQDILDWIYHTQGY 120

QY 230 FPDQNYTPGQVRYPLTFGCYKLVPEPKVEANKGENTSLHPVSLHGMDPPERV 289

Db 121 FPDQNYTPGQVRYPLTFGCYKLVPEPKVEANKGENTSLHPVSLHGMDPPERV 179

QY 290 LEWRFSRLAFHHVARELHPEYFKNC 315

Db 180 LEWRFSRLAFHHVARELHPEYFKNC 205

#### RESULT 10

S24985

nef protein - human immunodeficiency virus type 1

C: Species: human immunodeficiency virus type 1, HIV-1

C: Date: 20-Feb-1995 #sequence\_revision 20-Feb-1995 #text\_change 20-Sep-1999

A: Accession: S24985

R: Harris, M.; Hislop, S.; Patilinas, P.; Neil, J.C.

submitted to the EMBL Data Library, November 1991

A: Description: In vivo derived HIV-1 NEF genes are heterogeneous and lack detectable nuc

A: Reference number: S24985

A: Accession: S24985

A: Status: preliminary

A: Molecule type: DNA

A: Residues: 1-204 <HAR>

A: Cross-references: EMBL: X63044; NID: g60123; PIDN: CAA44770.1; PID: g60124

C: Superfamily: AIDS nef protein

Query Match 45.2%; Score 1019; DB 2; Length 204;  
Best Local Similarity 89.8%; Pred. No. 5.1e-72;  
Matches 185; Conservative 10; Mismatches 9; Indels 2; Gaps 1;

QY 110 MGGKWSKSVVGVPTVPRMRRAEPAADGVGAASRDLEKKGALITSSNTAATNAACAWLEA 169

Db 1 MGGKWSKSVVGVPTVPRMRRAEPAADGVGAASRDLEKKGALITSSNTAATNAACAWLEA 58

QY 170 QEEVEGFPVTPQVPLRPMTYKAAVDLSHFLKEKGLGLIHSQRQDILDWIYHTQGY 229

Db 59 QEEVEGFPVTPQVPLRPMTYKAAVDLSHFLKEKGLGLIHSQRQDILDWIYHTQGY 118

QY 230 FPDQNYTPGQVRYPLTFGCYKLVPEPKVEANKGENTSLHPVSLHGMDPPERV 289

Db 119 FPDQNYTPGQVRYPLTFGCYKLVPEPKVEANKGENTSLHPVSLHGMDPPERV 178

QY 290 LEWRFSRLAFHHVARELHPEYFKNC 315

Db 179 LEWRFSRLAFHHVARELHPEYFKNC 204

#### RESULT 11

ASLJBR

nef protein - human immunodeficiency virus type 1 (isolate BR)

N: Alternate names: 3'-orf protein; orf-F protein

C: Species: human immunodeficiency virus type 1, HIV-1

A: Note: host Homo sapiens (man)

C: Date: 31-Dec-1989 #sequence\_revision 31-Dec-1989 #text\_change 16-Jul-1999

A: Accession: D31667; S21993; S21995; S21997; S21999; S21991

R: Anand, R.; Thayer, R.; Srinivasan, A.; Nayyar, S.; Gardner, M.; Luciw, P.; Dandekar.

Virology 168, 79-89, 1989

A: Title: Biological and molecular characterization of human immunodeficiency virus (

A: Reference number: A94389; MUID: 89085613

A: Accession: D31667

A: Molecule type: DNA

A: Residues: 1-218 <ANA>

A: Cross-references: GB: M21098; NID: g326426; PIDN: AAA44222.1; PID: g326431

R: Steuler, H.; Storch-Hagenlocher, B.; Wildemann, B.; Hacke, W.

submitted to the EMBL Data Library, July 1991

A: Description: Distinct populations of HIV-1 in blood and cerebrospinal fluid as det

A: Reference number: S21990

A: Accession: S21993

A: Molecule type: DNA

A: Residues: 1-7 <STE>

A: Cross-references: EMBL: X61358; NID: g60177; PIDN: CAA43629.1; PID: g584027; EMBL: X613

A: Accession: S21995

A: Molecule type: DNA

A: Residues: 1-7 <ST2>

A: Cross-references: EMBL: X61355; NID: g60179; PIDN: CAA43623.1; PID: g584028; EMBL: X613

A: Accession: S21997

A: Molecule type: DNA

A: Residues: 1-7 <ST3>

A: Cross-references: EMBL: X61356; NID: g60181; PIDN: CAA43625.1; PID: g1129140; EMBL: X61

A: Accession: S21999

A: Molecule type: DNA

A: Residues: 1-7 <ST4>

A: Cross-references: EMBL: X61359; NID: g60182; PIDN: CAA43631.1; PID: g584030

C: Genetics:

A: Gene: nef; 3'-orf; orf-F

C: Superfamily: AIDS nef protein

C: Keywords: AIDS; immunodeficiency

Query Match 45.2%; Score 1019; DB 1; Length 218;

Best Local Similarity 84.1%; Pred. No. 5.5e-72;

Matches 185; Conservative 12; Mismatches 7; Indels 16; Gaps 2;

QY 110 MGGKWSKSVVGVPTVPRMR-----RAEPAADGVGAASRDLEKKGALITSS 155

Db 1 MGGKWSK--MAGWSTVPRMRRAEPARERMRRAEPAADGVGAASRDLEKKGALITSS 58

QY 156 NTAATNAACAWLEAQEEEEVEGFPVTPQVPLRPMTYKAAVDLSHFLKEKGLGLIHSQR 215

Db 59 NTAATNAACAWLEAQEEEEVEGFPVTPQVPLRPMTYKAAVDLSHFLKEKGLGLIHSQR 118

QY 216 QDILDWIYHTQGYPPDMQNYTPGCVRYPLTFGCYKLVPEPKVEANKGENTSLH 275

Db 119 QDILDWIYHTQGYPPDMQNYTPGCVRYPLTFGCYKLVPEPKVEANKGENTSLH 178

QY 276 PVSLSHGMDPPERLEWRFSRLAFHHVARELHPEYFKNC 315

Db 179 PVSLSHGMDPPERLEWRFSRLAFHHVARELHPEYFKNC 218

#### RESULT 12

S25937

nef protein - human immunodeficiency virus type 1

C: Species: human immunodeficiency virus type 1, HIV-1

C: Date: 25-Feb-1994 #sequence\_revision 01-Dec-1995 #text\_change 20-Sep-1999

C: Accession: S25937

R: Guo, H.G.; Chermann, J.C.; Waters, D.; Hall, L.; Louie, A.; Gallo, R.C.; Streicher

Nature 349, 745-746, 1991

A: Title: Sequence analysis of original HIV-1.

A:Reference number: S25937; MUID:91156044

A:Accession: S25937

A:Status: preliminary; nucleic acid sequence not shown; translation not shown

A:Molecule type: nucleic acid

A:Residues: 1-206 <GUO>

A:Cross-references: EMBL:X57465; NID:g60217; PIDN:CAA40702.1; PID:g60218

A:Note: the nucleotide sequence was submitted to the EMBL Data Library, March 1991

C:Superfamily: AIDS nef protein

Query Match 44.7%; Score 1007; DB 2; Length 206;

Best Local Similarity 85.9%; Pred. No. 4.4e-71;

Matches 177; Conservative 18; Mismatches 11; Indels 0; Gaps 0;

QY 110 MGGKWSKSSVVGWPTVRERMRRAEPAADGVGAASRDLEKHGAITSSNTAATNAACAWLEA 169

DB 1 MGGKWSKSSVVGWPTVRERMRRAEPAADGVGAASRDLEKHGAITSSNTAATNAACAWLEA 60

QY 170 QEEVEVGFPPVTPQVPLRPMTYKAAVDLSHFLKEKGGLEGLIHSORRQDILDWYHTQGY 229

DB 61 QEEVEVGFPPVTPQVPLRPMTYKAAVDLSHFLKEKGGLEGLIHSORRQDILDWYHTQGY 120

QY 230 FPDQNTYTPGVRYPYPLTFGWCYKLVPEPKVEANKGENTSLHPVSLHGMDDPEREV 289

DB 121 FPDQNTYTPGVRYPYPLTFGWCYKLVPEPKVEANKGENTSLHPVSLHGMDDPEREV 180

QY 290 LEWRFDSRLAFHHVARELHPYFKN 315

DB 181 LEWRFDSRLAFHHVARELHPYFKN 206

RESULT 13

ASLQ02

nef protein - human immunodeficiency virus type 1 (isolate ARV-2)

N:Alternate names: 3'-orf protein; Orf-F protein

C:Species: human immunodeficiency virus type 1, HIV-1

A:Note: host Homo sapiens (man)

C:Date: 17-May-1995 #sequence\_revision 17-May-1995 #text\_change 16-Jul-1999

C:Accession: A04009

R:Sanchez-Rescador, R.; Power, M.D.; Barr, P.J.; Stelmier, K.S.; Stempien, M.M.; Brown-SH

Science 227, 484-492, 1985

A:Title: Nucleotide sequence and expression of an AIDS-associated retrovirus (ARV-2).

A:Reference number: A04003; MUID:85090453

A:Accession: A04009

A:Molecule type: DNA

A:Residues: 1-210 <SAN>

A:Cross-references: GB:X02007; NID:g328658; PIDN:AAB59883.1; PID:g328667

C:Genetics:

A:Gene: nef; 3'-orf; orf-F

C:Superfamily: AIDS nef protein

C:Keywords: AIDS; immunodeficiency

Query Match 44.2%; Score 997; DB 1; Length 210;

Best Local Similarity 85.2%; Pred. No. 2.7e-70;

Matches 179; Conservative 16; Mismatches 11; Indels 4; Gaps 1;

QY 110 MGGKWSKSSVVGWPTVRERMR-----RAEPAADGVGAASRDLEKHGAITSSNTAATNAACA 165

DB 1 MGGKWSKSSVVGWPTVRERMRRAEPAADGVGAASRDLEKHGAITSSNTAATNAACA 60

QY 166 WLEAQEEVEVGFPPVTPQVPLRPMTYKAAVDLSHFLKEKGGLEGLIHSORRQDILDWYH 225

DB 61 WLEAQEEVEVGFPPVTPQVPLRPMTYKAAVDLSHFLKEKGGLEGLIHSORRQDILDWYH 120

QY 226 TQGYFPDQNTYTPGVRYPYPLTFGWCYKLVPEPKVEANKGENTSLHPVSLHGMDDP 285

DB 121 TQGYFPDQNTYTPGVRYPYPLTFGWCYKLVPEPKVEANKGENTSLHPVSLHGMDDP 180

QY 286 BREVLWRFDSRLAFHHVARELHPYFKN 315

DB 181 EKEVLWRFDSRLAFHHVARELHPYFKN 210

RESULT 14

I44001

nef protein - human immunodeficiency virus type 1 (strain YU-2)

N:Alternate names: 3'-orf protein; Orf-F protein

C:Species: human immunodeficiency virus type 1, HIV-1

A:Note: host Homo sapiens (man)

C:Date: 31-Dec-1993 #sequence\_revision 31-Dec-1993 #text\_change 23-Feb-1997

C:Accession: I44001

R:Li, Y.; Hui, H.; Burgess, C.J.; Price, R.W.; Sharp, P.M.; Hahn, B.H.; Shaw, G.M

J. Virol. 66, 6587-6600, 1992

A:Title: Complete nucleotide sequence, genome organization, and biological properties

A:Reference number: A44001; MUID:93021387

A:Accession: I44001

A:Molecule type: DNA

A:Residues: 1-214 <LIY>

A:Cross-references: GB:M93258

C:Genetics:

A:Gene: nef; 3'-orf; orf-F

C:Superfamily: AIDS nef protein

C:Keywords: AIDS; immunodeficiency

Query Match 44.0%; Score 991.5; DB 1; Length 214;

Best Local Similarity 84.2%; Pred. No. 7.5e-70;

Matches 181; Conservative 12; Mismatches 11; Indels 11; Gaps 2;

QY 110 MGGKWSKSSVVGWPTVR-----ERMRAEPAADGVGAASRDLEKHGAITSSNTAA 159

DB 1 MGGKWSKSSVVGWPTVRERMRRAEPAADGVGAASRDLEKHGAITSSNTAA 60

QY 160 TNAACAWLEAQEEVEVGFPPVTPQVPLRPMTYKAAVDLSHFLKEKGGLEGLIHSORRQDIL 219

DB 61 TNAACAWLEAQEEVEVGFPPVTPQVPLRPMTYKAAVDLSHFLKEKGGLEGLIHSORRQDIL 120

QY 220 DLWYHTQGYFPDQNTYTPGVRYPYPLTFGWCYKLVPEPKVEANKGENTSLHPVSL 279

DB 121 DLWYHTQGYFPDQNTYTPG-GTRWPLTFGWCYKLVPEPKVEANKGENTSLHPVSL 179

QY 280 HGMDPPERLEWRFDSRLAFHHVARELHPYFKN 314

DB 180 HGMDPPERLEWRFDSRLAFHHVARELHPYFKN 214

RESULT 15

S03247

nef protein (clone C15) - human immunodeficiency virus type 1 (fragment)

N:Alternate names: 3'-orf protein

C:Species: human immunodeficiency virus type 1, HIV-1

C:Date: 28-Feb-1990 #sequence\_revision 28-Feb-1990 #text\_change 20-Jun-2000

C:Accession: S03247

R:Ratner, L.; Starcich, B.; Josephs, S.F.; Hahn, B.H.; Reddy, E.P.; Livak, K.J.; Pett

Nucleic Acids Res. 13, 8219-8229, 1985

A:Title: Polymorphism of the 3' open reading frame of the virus associated with the a

A:Reference number: S03244; MUID:86067228

A:Accession: S03247

A:Molecule type: mRNA

A:Residues: 1-182 <RAT>

A:Cross-references: EMBL:X03189; NID:g61552; PIDN:CAA26948.1; PID:g1335562

C:Genetics:

A:Gene: nef; 3'-orf; orf-F

C:Superfamily: AIDS nef protein

Query Match 43.8%; Score 987; DB 2; Length 182;

Best Local Similarity 98.4%; Pred. No. 1.4e-69;

Matches 179; Conservative 1; Mismatches 2; Indels 0; Gaps 0;

QY 134 PAADGVGAASRDLEKHGAITSSNTAATNAACAWLEAQEEVEVGFPPVTPQVPLRPMTYKAA 193

DB 1 PAADGVGAASRDLEKHGAITSSNTAATNAACAWLEAQEEVEVGFPPVTPQVPLRPMTYKAA 60

QY 194 VDSLHFLKEKGGLEGLIHSORRQDILDWYHTQGYFPDQNTYTPGVRYPYPLTFGWCYK 253



Db 61 VDSLFLKEKGLSLHSQRQDILDWVHTQGYFPDQWNTYTPGVRPLTFGWCYK 120  
Qy 254 LVPVEPKVEANKGENTSLHVPVSLHGMDDPEREVLEWRDLSLAFHHVARELHPYFK 313  
Db 121 LVPVEPKVEANKGENTSLHVPVSLHGMDDPEREVLEWRDLSLAFHHVARELHPYFK 180  
Qy 314 NC 315  
Db 181 NC 182

RESULT 16  
S03246  
nef protein (clone HAT3) - human immunodeficiency virus type 1  
N:Alternate names: 3'-orf protein  
C:Species: human immunodeficiency virus type 1, HIV-1  
C:Date: 28-Feb-1990 #sequence\_revision 28-Feb-1990 #text\_change 20-Sep-1999  
C:Accession: S03246  
R:Ratner, L.; Starcich, B.; Josephs, S.F.; Hahn, B.H.; Reddy, E.P.; Livak, K.J.; Pettawa  
Nucleic Acids Res. 13, 8219-8229, 1985  
A:Title: Polymorphism of the 3' open reading frame of the virus associated with the acqu  
A:Reference number: S03244; MUID:8606728  
A:Accession: S03246  
A:Molecule type: DNA  
A:Residues: 1-204 <RAT>  
A:Cross-references: EMBL:X03190; NID:961550; PIDN:CAA26949.1; PID:g61551  
A:Note: The authors translated the codon AGT for residue 11 as Gly  
C:Genetics:  
A:Gene: nef; 3'-orf; Orf-F  
C:Superfamily: AIDS nef protein

Query Match 43.8%; Score 987; DB 2; Length 204;  
Best Local Similarity 84.5%; Pred. No. 1.6e-69;  
Matches 174; Conservative 16; Mismatches 14; Indels 2; Gaps 1;  
Qy 110 MGGKWSKSSVVGWPTVRERMPRAEPAADGVGAASRDLEKHGAISSNTAATNAACAWLEA 169  
Db 1 MGGKWSKSKMGWPAVRERMOKEAPADGVGAASRDLEKHGAISSNT--NNAACTWLEA 58  
Qy 170 QEEVEGFPVTPQVPLRPMTYKAAVDLSHFLKEKGLGLHSQRQDILDWVHTQGY 229  
Db 59 QEEVEGFPVTPQVPLRPMTYKAAVDLSHFLKEKGLGLHSQRQDILDWVHTQGY 118  
Qy 230 FPDQWNTYTPGVRPLTFGWCYKLVPEPKVEANKGENTSLHVPVSLHGMDDPEREV 289  
Db 119 FPDQWNTYTPGVRPLTFGWCYKLVPEPKVEANKGENTSLHVPVSLHGMDDPEREV 178  
Qy 290 LEWRDLSLAFHHVARELHPYFKNC 315  
Db 179 LVWKFDLSLAFHHVARELHPYFKNC 204

RESULT 17  
Q0LJZR  
nef protein - human immunodeficiency virus Zr-6  
N:Alternate names: 3'-orf protein; orf-F protein  
C:Species: human immunodeficiency virus Zr-6  
C:Date: 30-Sep-1987 #sequence\_revision 30-Sep-1987 #text\_change 16-Jul-1999  
C:Accession: F26192  
R:Srinivasan, A.; Anand, R.; York, D.; Ranganathan, P.; Feorino, P.; Schochetman, G.; Cu  
Gene 52, 71-82, 1987  
A:Title: Molecular characterization of human immunodeficiency virus from Zaire: nucleot  
A:Reference number: A26192; MUID:87248097  
A:Accession: F26192  
A:Molecule type: DNA  
A:Residues: 1-212 <SRI>  
A:Cross-references: GB:K03458; GB:M16322; NID:9329398; PIDN:AAA45381.1; PID:g329404  
C:Genetics:  
A:Gene: nef; 3'-orf; Orf-F  
C:Superfamily: AIDS nef protein  
C:Keywords: AIDS; immunodeficiency

Query Match 40.4%; Score 911; DB 1; Length 212;  
Best Local Similarity 77.4%; Pred. No. 1.4e-63;  
Matches 164; Conservative 23; Mismatches 19; Indels 6; Gaps 2;  
Qy 110 MGGKWSKSSVVGWPTVRERMPRAEPAADGVGAASRDLEKHGAISSNTAATNAAC 164  
Db 1 MGGKWSKSSVVGWPAVRERIRRTDPAADGVGAASRDLEKHGAISSNTDRTNADC 60  
Qy 165 AWLEAQEE--EEVGFPVTPQVPLRPMTYKAAVDLSHFLKEKGLGLHSQRQDILDW 223  
Db 61 AWLEAQEESEEVGFPVTPQVPLRPMTYKAAVDLSHFLKEKGLGLHSQRQDILDW 120  
Qy 224 YHTQGYFPDQWNTYTPGVRPLTFGWCYKLVPEPKVEANKGENTSLHVPVSLHGMDD 283  
Db 121 YHTQGYFPDQWNTYTPGVRPLTFGWCYKLVPEPKVEANKGENTSLHVPVSLHGMDD 180  
Qy 284 DPEREVLEWRDLSLAFHHVARELHPYFKNC 315  
Db 181 DPEREVLEWRDLSLAFHHVARELHPYFKNC 212  
RESULT 18  
Q0LJND  
nef protein - human immunodeficiency virus type 1 (isolate NDK)  
N:Alternate names: 3'-orf protein; orf-F protein  
C:Species: human immunodeficiency virus type 1, HIV-1  
A:Note: host Homo sapiens (man)  
C:Date: 30-Sep-1991 #sequence\_revision 30-Sep-1991 #text\_change 16-Jul-1999  
C:Accession: JQ0068  
R:Spire, B.; Sire, J.; Zachar, V.; Rey, F.; Barre-Sinoussi, F.; Galibert, F.; Hampe,  
Gene 81, 275-284, 1989  
A:Title: Nucleotide sequence of HIV1-NDK: a highly cytopathic strain of the human im  
A:Reference number: JQ0065; MUID:90034200  
A:Accession: JQ0068  
A:Molecule type: DNA  
A:Residues: 1-207 <SPT>  
A:Cross-references: GB:M27323; NID:g328154; PIDN:AAA44874.1; PID:g328163  
C:Genetics:  
A:Gene: nef  
C:Superfamily: AIDS nef protein  
C:Keywords: AIDS; immunodeficiency

Query Match 40.1%; Score 903.5; DB 1; Length 207;  
Best Local Similarity 77.3%; Pred. No. 5.1e-63;  
Matches 160; Conservative 25; Mismatches 21; Indels 1; Gaps 1;  
Qy 110 MGGKWSKSSVVGWPTVRERMPRAEPAADGVGAASRDLEKHGAISSNTAATNAACAWLEA 169  
Db 1 MGGKWSKSSVVGWPAVRERIRRTDPAADGVGAASRDLEKHGAISSNTASTNDTCWLEA 60  
Qy 170 QEE--EEVGFPVTPQVPLRPMTYKAAVDLSHFLKEKGLGLHSQRQDILDWVHTQGY 228  
Db 61 QEESEEVGFPVTPQVPLRPMTYKAAVDLSHFLKEKGLGLHSQRQDILDWVHTQGY 120  
Qy 229 YHTQGYFPDQWNTYTPGVRPLTFGWCYKLVPEPKVEANKGENTSLHVPVSLHGMDDPERE 288  
Db 121 YHTQGYFPDQWNTYTPGVRPLTFGWCYKLVPEPKVEANKGENTSLHVPVSLHGMDDPERE 180  
Qy 289 VLEWRDLSLAFHHVARELHPYFKNC 315  
Db 181 VLMWRDLSLAFHHVARELHPYFKNC 207  
RESULT 19  
B44963  
nef protein - human immunodeficiency virus type 1 (isolate Z321)  
N:Alternate names: 3'-orf protein; orf-F protein  
C:Species: human immunodeficiency virus type 1, HIV-1  
A:Note: host Homo sapiens (man)  
C:Date: 17-Feb-1994 #sequence\_revision 17-Feb-1994 #text\_change 28-May-1999

C:Accession: B44963  
R:Srinivasan, A.; York, D.; Butler Jr., D.; Jannoun-Nasr, R.; Getchell, J.; McCormick, J.  
AIDS Res. Hum. Retroviruses 5, 121-129, 1989  
A:Title: Molecular characterization of HIV-1 isolated from a serum collected in 1976: nu  
A:Reference number: A44963; MUID:89228766  
A:Accession: B44963  
A:Molecule type: DNA  
A:Residues: 1-205 <SRI>  
A:Cross-references: GB:M15896; NID:g329392; PIDN:AAB53951.1; PID:g329397  
C:Genetics:  
A:Gene: nef; 3'-orf; orf-F  
C:Superfamily: AIDS nef protein  
C:Keywords: AIDS; immunodeficiency

Query Match 37.1%; Score 837; DB 1; Length 205;  
Best Local Similarity 73.2%; Pred. No. 7.6e-58;  
Matches 153; Conservative 23; Mismatches 25; Indels 8; Gaps 3;

Qy 110 MGGKSKSVVGVPTVRMRRAEPAADGVGAASRDLEKHGAITSNTAATNAACAW 166

Db 1 MGNKWSK---GWPVAVRIRIOTPPAPAAEGVGAASQDLAKHGAISSNTATNPNPCAW 56

Qy 167 LEAQEE-EVGFPTVPTQVPLRPMYKAAVDLSHFLKKGGLGLIHSORQDILDWIYH 225

Db 57 LEAQEESEVGVPTVPTQVPLRPMYKAAVDLSHFLKKGGLGLIHSORQDILDWIYH 116

Qy 226 TQGYFPDQNTYTPGCVRYPLTFGWCYKLVPEPKVEANKGENTSLHVPVSLHGMDDP 285

Db 117 TQGYFPDQNTYTPGCVRYPLTFGWCYKLVPEPKVEANKGENTSLHVPVSLHGMDDP 176

Qy 286 EREVLEWFDLSRLAFHHVARELHPEYFKN 314

Db 177 EREVLEWFDLSRLAFHHVARELHPEYFKN 205

## RESULT 20

T01673

nef protein - human immunodeficiency virus type 1

C:Species: human immunodeficiency virus type 1, HIV-1

C&gt;Date: 19-Feb-1999 #sequence\_revision 19-Feb-1999 #text\_change 21-Jul-2000

C:Accession: T01673

R:Alizon, M.; Wain-Hobson, S.; Gluckman, J.C.; Sonigo, P.

Cell 46, 63-74, 1986

A:Title: Genetic variability of the AIDS virus: Nucleotide sequence analysis of two isol

A:Reference number: 214389; MUID:86245056

A:Accession: T01673

A:Status: preliminary; translated from GB/EMBL/DDDBJ

A:Molecule type: mRNA

A:Residues: 1-209 &lt;ALI&gt;

A:Cross-references: EMBL:K03456; NID:g60228; PIDN:CAA28017.1; PID:g60235

C:Superfamily: AIDS nef protein

Query Match 37.0%; Score 834.5; DB 2; Length 209;  
Best Local Similarity 70.1%; Pred. No. 1.2e-57;  
Matches 148; Conservative 26; Mismatches 30; Indels 7; Gaps 2;

Qy 110 MGGKSKSVVGVPTVRMRRAEPAADGVG----AASRDLEKHGAITSNTAATNAAC 164

Db 1 MGGKSKSVVGVPTVRMRRAEPAADGVG----AASRDLEKHGAITSNTAATNAAC 60

Qy 165 AWLEAQEEVGFPTVPTQVPLRPMYKAAVDLSHFLKKGGLGLIHSORQDILDWIY 224

Db 61 E--PPEEEVGFPTVPTQVPLRPMYKAAVDLSHFLKKGGLGLIHSORQDILDWIY 118

Qy 225 HQGYFPDQNTYTPGCVRYPLTFGWCYKLVPEPKVEANKGENTSLHVPVSLHGMDD 284

Db 119 HQGYFPDQNTYTPGCVRYPLTFGWCYKLVPEPKVEANKGENTSLHVPVSLHGMDD 178

Qy 285 EREVLEWFDLSRLAFHHVARELHPEYFKN 315

Db 179 EREVLEWFDLSRLAFHHVARELHPEYFKN 209

## RESULT 21

ASLJK

nef protein - simian immunodeficiency virus SIVcpz

N:Alternate names: 3'-orf protein; orf-F protein

C:Species: simian immunodeficiency virus SIVcpz

A:Note: host Pan troglodytes (chimpanzee)

C&gt;Date: 30-Sep-1991 #sequence\_revision 30-Sep-1991 #text\_change 16-Jul-1999

C:Accession: S09991

R:Huet, T.; Cheynier, R.; Meyerhans, A.; Roelants, G.; Wain-Hobson, S.

Nature 345, 356-359, 1990

A:Title: Genetic organization of a chimpanzee lentivirus related to HIV-1.

A:Reference number: S09983; MUID:90259077

A:Accession: S09991

A:Status: nucleic acid sequence not shown; translation not shown

A:Molecule type: DNA

A:Residues: 1-205 &lt;HUE&gt;

A:Cross-references: EMBL:X52154; NID:g58866; PIDN:CAA36408.1; PID:g58877

C:Genetics:

A:Gene: nef

C:Superfamily: AIDS nef protein

C:Keywords: AIDS; immunodeficiency

Query Match 36.9%; Score 832; DB 1; Length 205;  
Best Local Similarity 70.2%; Pred. No. 1.9e-57;  
Matches 144; Conservative 27; Mismatches 34; Indels 0; Gaps 0;

Qy 110 MGGKSKSVVGVPTVRMRRAEPAADGVGAASRDLEKHGAITSNTAATNAACAWLEA 169

Db 1 MGTKWSKSLVGVPEVRRIRREAPTAAGVGVSKDLRHGAITSNTPETNQTALWLEE 60

Qy 170 QEEVEVGFPTVPTQVPLRPMYKAAVDLSHFLKKGGLGLIHSORQDILDWIYHTQGY 229

Db 61 MDNEVGFPTVPTQVPLRPMYKAAVDLSHFLKKGGLGLIHSORQDILDWIYHTQGF 120

Qy 230 FPDQNTYTPGCVRYPLTFGWCYKLVPEPKVEANKGENTSLHVPVSLHGMDDPDERV 289

Db 121 FPDQNTYTPGCVRYPLTFGWCYKLVPEPKVEANKGENTSLHVPVSLHGMDDPDERV 180

Qy 290 LEWRFDSRLAFHHVARELHPEYFKN 314

Db 181 LVWRFDSRLAFHHVARELHPEYFKN 205

## RESULT 22

S59931

glycerophosphodiester phosphodiesterase (EC 3.1.4.46) - Haemophilus influenzae (strai

N:Alternate names: Igd-binding protein; protein D

C:Species: Haemophilus influenzae

A:Variety: strain 3639

C&gt;Date: 15-Feb-1996 #sequence\_revision 01-Mar-1996 #text\_change 08-Oct-1999

C:Accession: S59931; S47333

R:Song, X.M.; Forsgren, A.; Janson, H.

Infect. Immun. 63, 696-699, 1995

A:Title: The gene encoding protein D (hpd) is highly conserved among Haemophilus infl

A:Reference number: S59931; MUID:95122210

A:Accession: S59931

A:Status: preliminary

A:Molecule type: DNA

A:Residues: 1-364 &lt;SON&gt;

A:Cross-references: EMBL:Z35656; NID:g525213; PIDN:CAA84715.1; PID:g525214

A:Experimental source: strain 3639

C:Keywords: lipoprotein; phosphoric diester hydrolase; surface antigen

Query Match 25.7%; Score 578.5; DB 2; Length 364;  
Best Local Similarity 47.1%; Pred. No. 2.1e-37;  
Matches 145; Conservative 19; Mismatches 71; Indels 73; Gaps 11;

Qy 1 CSSHSSNMANTQMSKIIIAHRGAGYLPETLESKALFAQQADYLSODLAWTKDGLR 60

Db 1 CSSHSSNMANTQMSKIIIAHRGAGYLPETLESKALFAQQADYLSODLAWTKDGLR 111

Db 19 CSSHSSNNANTOMKSDKIIIAHRCASGYLPEHTLESKALAFQAQADYLEQDLAMTKDGR 78  
Qy 61 VVIHDFLDGLTDVAKKPPHRRKDGRIYVDFTLKEIQSLEMTENFETMGK----- 113  
Db 79 VVIHDFLDGLTDVAKKPPHRRKDGRIYVDFTLKEIQSLEMTENFETMGKQAQVYPN 138  
Qy 114 ----W-----SKSSVVG-WPTVRRMRRAEPAADGVGAASRDLE 147  
Db 139 RFLWKSHFRHTFEDEIEFTQGLEKSTGKVGIIPEIKAPWFHHQNGKDIAAETLKV 198  
Qy 148 KHGAITSSNTAATNAACAWLEAEVEEFGFVPTQVPLRPMTYKAAVDLSHFLKEKGGLE 207  
Db 199 KYGDKKTD-----MVYLQTFDFNELKRIKTELLPQMGMDLK-LVQLIAYTDWK---- 246  
Qy 208 GLIHSRRQDILDWIYHTQGYFPDQW-----NYTPGPGVRYPLTFGWCYKLV 257  
Db 247 ----ETQEKDPKGYW----NYNDWMFKPGAMAEVVKYADGVGP-----GW-YMLV 291  
Qy 258 E---PDKV 262  
Db 292 EESKPDNI 299

Query Match 25.7%; Score 578.5; DB 2; Length 364;  
Best Local Similarity 47.1%; Pred. No. 2.1e-37;  
Matches 145; Conservative 19; Mismatches 71; Indels 73; Gaps 11;  
RESULT 23  
S59932  
glycerophosphodiester phosphodiesterase (EC 3.1.4.46) - Haemophilus influenzae (strain 3  
N:Alternate names: immunoglobulin D-binding protein; protein D  
C:Species: Haemophilus influenzae  
A:Variety: strain 3640  
C:Date: 15-Feb-1996 #sequence\_revision 01-Mar-1996 #text\_change 08-Oct-1999  
C:Accession: S59932; S47334  
R:Song, X.M.; Forsgren, A.; Janson, H.  
Infect. Immun. 63, 696-699, 1995  
A:Title: The gene encoding protein D (hpd) is highly conserved among Haemophilus influen  
A:Reference number: S59931; MUID:95122210  
A:Accession: S59932  
A:Molecule type: DNA  
A:Residues: 1-364 <SON>  
A:Cross-references: EMBL:Z35657; NID:9525215; PIDN:CAA84716.1; PID:9525216  
A:Experimental source: strain 3640  
C:Keywords: lipoprotein; phosphoric diester hydrolase; surface antigen

Query Match 25.7%; Score 578.5; DB 2; Length 364;  
Best Local Similarity 47.1%; Pred. No. 2.1e-37;  
Matches 145; Conservative 19; Mismatches 71; Indels 73; Gaps 11;

Qy 1 CSSHSSNNANTOMKSDKIIIAHRCASGYLPEHTLESKALAFQAQADYLEQDLAMTKDGR 60  
Db 19 CSSHSSNNANTOMKSDKIIIAHRCASGYLPEHTLESKALAFQAQADYLEQDLAMTKDGR 78  
Qy 61 VVIHDFLDGLTDVAKKPPHRRKDGRIYVDFTLKEIQSLEMTENFETMGK----- 113  
Db 79 VVIHDFLDGLTDVAKKPPHRRKDGRIYVDFTLKEIQSLEMTENFETMGKQAQVYPN 138  
Qy 114 ----W-----SKSSVVG-WPTVRRMRRAEPAADGVGAASRDLE 147  
Db 139 RFLWKSHFRHTFEDEIEFTQGLEKSTGKVGIIPEIKAPWFHHQNGKDIAAETLKV 198  
Qy 148 KHGAITSSNTAATNAACAWLEAEVEEFGFVPTQVPLRPMTYKAAVDLSHFLKEKGGLE 207  
Db 199 KYGDKKTD-----MVYLQTFDFNELKRIKTELLPQMGMDLK-LVQLIAYTDWK---- 246  
Qy 208 GLIHSRRQDILDWIYHTQGYFPDQW-----NYTPGPGVRYPLTFGWCYKLV 257  
Db 247 ----ETQEKDPKGYW----NYNDWMFKPGAMAEVVKYADGVGP-----GW-YMLV 291  
Qy 258 E---PDKV 262  
Db 292 EESKPDNI 299

RESULT 24

A43576  
glycerophosphodiester phosphodiesterase (EC 3.1.4.46) precursor - Haemophilus influe  
N:Alternate names: immunoglobulin D-binding protein; protein D  
C:Species: Haemophilus influenzae  
A:Variety: isolate 772  
C:Date: 22-Jan-1993 #sequence\_revision 22-Jan-1993 #text\_change 08-Oct-1999  
C:Accession: A43576  
R:Janson, H.; Heden, L.O.; Grubb, A.; Ruan, M.; Forsgren, A.  
Infect. Immun. 59, 119-125, 1991  
A:Title: Protein D, an immunoglobulin D-binding protein of Haemophilus influenzae: c  
A:Reference number: A43576; MUID:91099948  
A:Accession: A43576  
A:Molecule type: DNA  
A:Residues: 1-364 <JAN>  
A:Cross-references: GB:M37487; NID:9148970; PIDN:AAA24998.1; PID:9148971  
A:Experimental source: isolate 772  
C:Keywords: lipoprotein; phosphoric diester hydrolase; surface antigen  
F:1-18/Domain: signal sequence #status predicted <SIG>  
F:19-364/Product: glycerophosphodiester phosphodiesterase #status predicted <MAT>

Query Match 25.7%; Score 578.5; DB 2; Length 364;  
Best Local Similarity 47.1%; Pred. No. 2.1e-37;  
Matches 145; Conservative 19; Mismatches 71; Indels 73; Gaps 11;  
Qy 1 CSSHSSNNANTOMKSDKIIIAHRCASGYLPEHTLESKALAFQAQADYLEQDLAMTKDGR 60  
Db 19 CSSHSSNNANTOMKSDKIIIAHRCASGYLPEHTLESKALAFQAQADYLEQDLAMTKDGR 78  
Qy 61 VVIHDFLDGLTDVAKKPPHRRKDGRIYVDFTLKEIQSLEMTENFETMGK----- 113  
Db 79 VVIHDFLDGLTDVAKKPPHRRKDGRIYVDFTLKEIQSLEMTENFETMGKQAQVYPN 138  
Qy 114 ----W-----SKSSVVG-WPTVRRMRRAEPAADGVGAASRDLE 147  
Db 139 RFLWKSHFRHTFEDEIEFTQGLEKSTGKVGIIPEIKAPWFHHQNGKDIAAETLKV 198  
Qy 148 KHGAITSSNTAATNAACAWLEAEVEEFGFVPTQVPLRPMTYKAAVDLSHFLKEKGGLE 207  
Db 199 KYGDKKTD-----MVYLQTFDFNELKRIKTELLPQMGMDLK-LVQLIAYTDWK---- 246  
Qy 208 GLIHSRRQDILDWIYHTQGYFPDQW-----NYTPGPGVRYPLTFGWCYKLV 257  
Db 247 ----ETQEKDPKGYW----NYNDWMFKPGAMAEVVKYADGVGP-----GW-YMLV 291  
Qy 258 E---PDKV 262  
Db 292 EESKPDNI 299

RESULT 25  
S59934  
glycerophosphodiester phosphodiesterase (EC 3.1.4.46) precursor - Haemophilus influe  
N:Alternate names: immunoglobulin D-binding protein; protein D  
C:Species: Haemophilus influenzae  
A:Variety: strain Eagan; strain HK695  
C:Date: 15-Feb-1996 #sequence\_revision 01-Mar-1996 #text\_change 08-Oct-1999  
C:Accession: S59934; S59935; S47336; S47337  
R:Song, X.M.; Forsgren, A.; Janson, H.  
Infect. Immun. 63, 696-699, 1995  
A:Title: The gene encoding protein D (hpd) is highly conserved among Haemophilus inf  
A:Reference number: S59931; MUID:95122210  
A:Accession: S59934  
A:Molecule type: DNA  
A:Residues: 1-364 <SON>  
A:Cross-references: EMBL:Z35659; NID:9525219; PIDN:CAA84718.1; PID:9525220  
A:Experimental source: strain Eagan  
A:Accession: S59935  
A:Molecule type: DNA  
A:Residues: 1-364 <SON>  
A:Cross-references: EMBL:Z35660; NID:9525221; PIDN:CAA84719.1; PID:9525222  
A:Experimental source: strain HK695  
C:Keywords: lipoprotein; phosphoric diester hydrolase; surface antigen

F:1-18/Domain: signal sequence #status predicted <SIG>  
F:19-364/Product: protein D #status predicted <NAT>

Query Match 25.7%; Score 578.5; DB 2; Length 364;  
Best Local Similarity 47.1%; Pred. No. 2.1e-37;  
Matches 145; Conservative 19; Mismatches 71; Indels 73; Gaps 11;

QY 1 CSSHSSNNANTOMKSDKIIIAHAGSGYLPHTLESKALAFQAQADYLEQDLAMTKDGR 60  
DB 19 CSSHSSNNANTOMKSDKIIIAHAGSGYLPHTLESKALAFQAQADYLEQDLAMTKDGR 78  
QY 61 VVIHDFLDGLTDVAKKPPHRRKGRYVYIDFTLKEIQSLEMTENFETMGK----- 113  
DB 79 VVIHDFLDGLTDVAKKPPHRRKGRYVYIDFTLKEIQSLEMTENFETMGKQAQVYPN 138  
QY 114 ----W-----SKSSVVG-WPTVRMRRAEPAADGVGAASRDLE 147  
DB 139 RPLWKSHFRHTFEDEIEFIQGLEKSTGKKGVIPEIKAPWFHONGKDIAETLKV 198  
QY 148 KHGATSSNTAATNAACAWLEAQEEVEGFPVTPQVPLRPMTYKAAVDLSHFLKEKG 207  
DB 199 KYGYDKKTD-----MVYLQTFDFNELKRIKTELLPQMGMDLK-LVQLIAYTD 246  
QY 208 GLIHSQRQDILDLIWYHTQGYFPDQW-----NYTPGPGVRYPLTFGWCYK 257  
DB 247 ----ETQEKDPKGYW-----NYNDMMFKPGAMAEVVKYADGVG-----GW-YMLV 291  
QY 258 E---PDKV 262  
DB 292 EESKPDNI 299

RESULT 26  
S59933  
glycerophosphodiester phosphodiesterase (EC 3.1.4.46) - Haemophilus influenzae (strain 6  
N:Alternate names: immunoglobulin D-binding protein; protein D  
C:Species: Haemophilus influenzae  
A:Variety: strain 6-7626  
C:Date: 15-Feb-1996 #sequence\_revision 01-Mar-1996 #text\_change 08-Oct-1999  
C:Accession: S59933; S47335  
R:Song, X.M.; Forsgren, A.; Janson, H.  
Infect. Immun. 63, 696-699, 1995  
A:Title: The gene encoding protein D (hpd) is highly conserved among Haemophilus influen  
A:Reference number: S59931; MUID:95122210  
A:Accession: S59933  
A:Molecule type: DNA  
A:Residues: 1-364 <SON>  
A:Cross-references: EMBL:235658; NID:9525217; PIDN:CAAB4717.1; PID:9525218  
A:Experimental source: strain 6-7626  
C:Keywords: lipoprotein; phosphoric diester hydrolase; surface antigen

Query Match 25.4%; Score 572.5; DB 2; Length 364;  
Best Local Similarity 46.8%; Pred. No. 6.1e-37;  
Matches 144; Conservative 19; Mismatches 77; Indels 73; Gaps 11;

QY 1 CSSHSSNNANTOMKSDKIIIAHAGSGYLPHTLESKALAFQAQADYLEQDLAMTKDGR 60  
DB 19 CSSHSSNNANTOMKSDKIIIAHAGSGYLPHTLESKALAFQAQADYLEQDLAMTKDGR 78  
QY 61 VVIHDFLDGLTDVAKKPPHRRKGRYVYIDFTLKEIQSLEMTENFETMGK----- 113  
DB 79 VVIHDFLDGLTDVAKKPPHRRKGRYVYIDFTLKEIQSLEMTENFETMGKQAQVYPN 138  
QY 114 ----W-----SKSSVVG-WPTVRMRRAEPAADGVGAASRDLE 147  
DB 139 RPLWKSHFRHTFEDEIEFIQGLEKSTGKKGVIPEIKAPWFHONGKDIAETLKV 198  
QY 148 KHGATSSNTAATNAACAWLEAQEEVEGFPVTPQVPLRPMTYKAAVDLSHFLKEKG 207  
DB 199 KYGYDKKTD-----MVYLQTFDFNELKRIKTELLPQMGMDLK-LVQLIAYTD 246

QY 208 GLIHSQRQDILDLIWYHTQGYFPDQW-----NYTPGPGVRYPLTFGWCYK 257  
DB 247 ----ETQEKDPKGYW-----NYNDMMFKPGAMAEVVKYADGVG-----GW-YMLV 291  
QY 258 E---PDKV 262  
DB 292 EESKPDNI 299

## RESULT 27

G64086  
glycerophosphodiester phosphodiesterase (EC 3.1.4.46) - Haemophilus influenzae (stra.  
N:Alternate names: Igd-binding protein; protein D  
C:Species: Haemophilus influenzae  
C:Date: 18-Aug-1995 #sequence\_revision 18-Aug-1995 #text\_change 21-Jul-2000  
C:Accession: G64086  
R:Fleischmann, R.D.; Adams, M.D.; White, O.; Clayton, R.A.; Kirkness, E.F.; Kerlavage  
; Godayne, J.D.; Scott, J.; Shirley, R.; Liu, L.I.; Glodek, A.; Kelley, J.M.; Weidman  
; D.M.; Brandon, R.C.; Fine, L.D.; Fritchman, J.L.; Fuhrmann, J.L.; Geoghegan, N.S.M.  
Science 269, 496-512, 1995  
A:Authors: Gnehm, C.L.; McDonald, L.A.; Small, K.V.; Fraser, C.M.; Smith, H.O.; Vente  
A:Title: Whole-genome random sequencing and assembly of Haemophilus influenzae Rd.  
A:Reference number: A64000; MUID:95350630  
A:Accession: G64086  
A:Status: nucleic acid sequence not shown; translation not shown  
A:Molecule type: DNA  
A:Residues: 1-364 <TIGR>  
A:Cross-references: GB:U32751; GB:L42023; NID:93212200; PIDN:AAC22348.1; PID:gl573690  
A:Experimental source: strain Rd KW20  
C:Keywords: lipoprotein; phosphoric diester hydrolase; surface antigen

Query Match 24.9%; Score 562.5; DB 2; Length 364;  
Best Local Similarity 46.1%; Pred. No. 3.7e-36;  
Matches 142; Conservative 21; Mismatches 72; Indels 73; Gaps 11;

QY 1 CSSHSSNNANTOMKSDKIIIAHAGSGYLPHTLESKALAFQAQADYLEQDLAMTKDGR 60  
DB 19 CSSHSSNNANTOMKSDKIIIAHAGSGYLPHTLESKALAFQAQADYLEQDLAMTKDGR 78  
QY 61 VVIHDFLDGLTDVAKKPPHRRKGRYVYIDFTLKEIQSLEMTENFETMGK----- 113  
DB 79 VVIHDFLDGLTDVAKKPPHRRKGRYVYIDFTLKEIQSLEMTENFETMGKQAQVYPN 138  
QY 114 ----W-----SKSSVVG-WPTVRMRRAEPAADGVGAASRDLE 147  
DB 139 RPLWKSHFRHTFEDEIEFIQGLEKSTGKKGVIPEIKAPWFHONGKDIAETLKV 198  
QY 148 KHGATSSNTAATNAACAWLEAQEEVEGFPVTPQVPLRPMTYKAAVDLSHFLKEKG 207  
DB 199 KYGYDKKTD-----MVYLQTFDFNELKRIKTELLPQMGMDLK-LVQLIAYTD 246  
QY 208 GLIHSQRQDILDLIWYHTQGYFPDQW-----NYTPGPGVRYPLTFGWCYK 257  
DB 247 ----ETQEKDPKGYW-----NYNDMMFKPGAMAEVVKYADGVG-----GW-YMLV 291  
QY 258 E---PDKV 262  
DB 292 EESKPDNI 299

## RESULT 28

S59936  
glycerophosphodiester phosphodiesterase (EC 3.1.4.46) - Haemophilus influenzae (stra  
N:Alternate names: Igd-binding protein; protein D  
C:Species: Haemophilus influenzae  
A:Variety: strain NCTC 8468  
C:Date: 15-Feb-1996 #sequence\_revision 01-Mar-1996 #text\_change 08-Oct-1999  
C:Accession: S59936; S47338  
R:Song, X.M.; Forsgren, A.; Janson, H.  
Infect. Immun. 63, 696-699, 1995  
A:Title: The gene encoding protein D (hpd) is highly conserved among Haemophilus infl  
A:Reference number: S59931; MUID:95122210

A:Accession: S59936  
A:Molecule type: DNA  
A:Residues: 1-364 <SON>  
A:Cross-references: EMBL:235661; NID:g525223; PIDN:CAA84720.1; PID:g525224  
A:Experimental source: strain NCTC 8468  
C:Keywords: lipoprotein; phosphoric diester hydrolase; surface antigen

Query Match 24.78; Score 556.5; DB 2; Length 364;  
Best Local Similarity 45.8%; Pred. No. 1.1e-35;  
Matches 141; Conservative 20; Mismatches 74; Indels 73; Gaps 11;

QY 1 CSSHSSNANTQMSDKIIIAHAGSGYLPETHLESKALFAQADYLEODLAMTKDGLR 60  
DB 19 CSSHSSNANTQMSDKIIIAHAGSGYLPETHLESKALFAQADYLEODLAMTKDGLR 78  
QY 61 VVHDFHDLGLTDVAKKPPHRRKGRYYVIDFTLKEIOSLEMTENFTTMGK----- 113  
DB 79 VVHDFHDLGLTDVAKKPPHRRKGRYYVIDFTLKEIOSLEMTENFTTMGKQAQVYPN 138  
QY 114 ---W-----SKSSVVG-WPTVRRMRRAEPAADGVAASRDLE 147  
DB 139 RPPLWKSHFRHTFDETEFTQGLEKSTGKKVGIYPEIKAPWFHHONGKDAIATLKLK 198  
QY 148 KHGAITSSNTAATNAACAALEAEQEEVFPVTPQVPLRPMYKAAVDLSHLFKEKGGLE 207  
DB 199 KYGDKD-----MYVLQTFDFNELKRIKTELLPQMGHDLK-LVOLIAYTDWK----- 246  
QY 208 GLIHSQRQDILDLYIHTQGYFPDQ-----NYTPGPGVRYPLTGTWCYKLVVPV 257  
DB 247 ----ETQEKDPKGYWV----NYNYDMFKPGAMAEVWYADGVP-----GW-YMLVYNK 291  
QY 258 E---PDKV 262  
DB 292 EESKPDNI 299

## RESULT 29

S33982  
trans-activating transcription regulator - human immunodeficiency virus type 1  
C:Species: human immunodeficiency virus type 1, HIV-1  
C:Date: 06-Oct-1994 #sequence\_revision 23-Feb-1996 #text\_change 20-Sep-1999  
C:Accession: S33982; S26385; S19864

submitted to the EMBL Data Library, November 1991

A:Reference number: S33979

A:Accession: S33982

A:Molecule type: DNA

A:Residues: 1-86 <CAR>

A:Cross-references: EMBL:Z11530; NID:g60192; PIDN:CAA77625.1; PID:g60196

R:Siderovski, D.P.; Matsuyama, T.; Frigerio, E.; Chui, S.; Min, X.; Erfle, H.; Summer-Sm

Nucleic Acids Res. 20, 5311-5320, 1992

A:Title: Random mutagenesis of the human immunodeficiency virus type-1 trans-activator c

A:Reference number: S26385; MUID:93065196

A:Accession: S26385

A:Molecule type: nucleic acid

A:Residues: 1-86 <SID>

A:Cross-references: EMBL:X64650; NID:g60144; PIDN:CAA45921.1; PID:g60145

C:Genetics:

A:Gene: tat

A:Introns: 72/2

C:Superfamily: AIDS trans-activating transcription regulator

C:Keywords: AIDS; immunodeficiency

Query Match 21.6%; Score 486; DB 2; Length 86;  
Best Local Similarity 100.0%; Pred. No. 5.6e-31;  
Matches 85; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 318 EPVDPRLPEPWKHPGSPQKTACTNCYCKKCCFHCQVCFTTKALGISYGRKKRRRRPQ 377  
DB 2 EPVDPRLPEPWKHPGSPQKTACTNCYCKKCCFHCQVCFTTKALGISYGRKKRRRRPQ 61

QY 378 SOTHQVSLSKOPTSQSRGDPGTGPKE 402  
DB 62 SOTHQVSLSKOPTSQSRGDPGTGPKE 86

## RESULT 30

TNLJ12

trans-activating transcription regulator - human immunodeficiency virus type 1 (isol  
C:Species: human immunodeficiency virus type 1, HIV-1  
A:Note: host Homo sapiens (man)

C:Date: 04-Dec-1986 #sequence\_revision 04-Dec-1986 #text\_change 02-Jul-1998

C:Accession: A04017

R:Arva, S.K.; Gallo, R.C.

Proc. Natl. Acad. Sci. U.S.A. 83, 2209-2213, 1986

A:Title: Three novel genes of human T-lymphotropic virus type III: immune reactivity

A:Reference number: A94093; MUID:86177573

A:Accession: A04017

A:Molecule type: DNA

A:Residues: 1-95 <ARY>

C:Genetics:

A:Gene: tat

C:Superfamily: AIDS trans-activating transcription regulator

C:Keywords: AIDS; immunodeficiency; transcription regulation

Query Match 21.2%; Score 478; DB 1; Length 95;

Best Local Similarity 98.8%; Pred. No. 2.7e-30;

Matches 84; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 318 EPVDPRLPEPWKHPGSPQKTACTNCYCKKCCFHCQVCFTTKALGISYGRKKRRRRPQ 377

DB 11 EPVDPRLPEPWKHPGSPQKTACTNCYCKKCCFHCQVCFTTKALGISYGRKKRRRRPQ 70

QY 378 SOTHQVSLSKOPTSQSRGDPGTGPKE 402

DB 71 SOTHQVSLSKOPTSQSRGDPGTGPKE 95

Search completed: August 26, 2002, 08:13:22

Job time: 291 sec

GenCore version 4.5  
Copyright (c) 1993 - 2000 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: August 26, 2002, 08:15:41 ; Search time 40.39 Seconds  
(without alignments)  
394.002 Million cell updates/sec

Title: US-09-509-239-17

Perfect score: 2255

Sequence: 1 CSSHSSNMANTQMKSDKIII.....QSRDPTGPKETSCHHHHHH 411

Scoring table:

BLOSUM62  
Gapop 10.0 , Gapext 0.5

Searched: 105224 seqs, 38719550 residues

Total number of hits satisfying chosen parameters: 105224

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : SwissProt\_40.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

# SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	1130	50.1	206	1 NEF_HV1BR	P03406 human immun
2	1114	49.4	206	1 NEF_HV1PV	P03405 human immun
3	1111	49.3	206	1 NEF_HV1W	O70627 human immun
4	1108	49.1	206	1 NEF_HV112	P04324 human immun
5	1088.5	48.3	205	1 NEF_HV1B8	P05855 human immun
6	1019	45.2	218	1 NEF_HV1B8	P12479 human immun
7	1005.5	44.6	205	1 NEF_HV1B3	P19545 human immun
8	997.5	44.2	211	1 NEF_HV1OY	P20886 human immun
9	997	44.2	210	1 NEF_HV1A2	P03407 human immun
10	994	44.1	208	1 NEF_HV1RH	P05858 human immun
11	991.5	44.0	214	1 NEF_HV1Y2	P35959 human immun
12	981	43.5	216	1 NEF_HV1JR	P20867 human immun
13	974	43.2	208	1 NEF_HV1S1	P19546 human immun
14	927.5	41.1	206	1 NEF_HV1EL	P04604 human immun
15	911	40.4	212	1 NEF_HV1Z6	P04602 human immun
16	903.5	40.1	207	1 NEF_HV1ND	P18801 human immun
17	872	38.7	205	1 NEF_HV1U4	P24741 human immun
18	842	37.3	182	1 NEF_HV1MN	P05856 human immun
19	837	37.1	205	1 NEF_HV1ZH	P05859 human immun
20	834.5	37.0	209	1 NEF_HV1MA	P04503 human immun
21	832	36.9	205	1 NEF_HV1CZ	P17664 chimpanzee
22	819	36.3	239	1 NEF_HV1SC	P05857 human immun
23	645	28.6	123	1 NEF_HV1H2	P04601 human immun
24	638	28.3	123	1 NEF_HV1B1	P03404 human immun
25	562.5	24.9	364	1 GLPQ_HAEIN	O06282 haemophilus
26	486	21.6	86	1 TAT_HV1B1	P04606 human immun
27	481	21.3	86	1 TAT_HV1PV	P04607 human immun
28	478	21.2	86	1 TAT_HV112	P04326 human immun
29	470	20.8	86	1 TAT_HV1BR	P04610 human immun
30	461	20.4	102	1 TAT_HV1JR	P05908 human immun
31	454	20.1	86	1 TAT_HV1H2	P04608 human immun
32	446	19.8	101	1 TAT_HV1JR	P20879 human immun
33	439.5	19.5	97	1 NEF_HV1Z2	P12478 human immun

34	439	19.5	253	1 NEF_HV2KR	O74127 human immun
35	437	19.4	101	1 TAT_HV1SC	P05906 human immun
36	435	19.3	101	1 TAT_HV1C4	P05907 human immun
37	435	19.3	101	1 TAT_HV1S1	P19553 human immun
38	434	19.2	101	1 TAT_HV1Y2	P35965 human immun
39	431	19.1	101	1 TAT_HV1MN	P05905 human immun
40	431	19.1	101	1 TAT_HV1S3	P19552 human immun
41	429.5	19.0	309	1 NEF_HV1S4	P12482 simian immun
42	426.5	18.9	261	1 NEF_HV1SVP	P19501 simian immun
43	421	18.7	257	1 NEF_HV2D1	P17753 human immun
44	419	18.6	101	1 TAT_HV1A2	P04614 human immun
45	419	18.6	101	1 TAT_HV1OY	P20893 human immun

## ALIGNMENTS

RESULT 1					
NEF_HV1BR					
ID	NEF_HV1BR	STANDARD:	PRT:	206 AA.	
AC	P03406;				
DT	21-JUL-1986 (Rel. 01, Created)				
DT	21-JUL-1986 (Rel. 01, Last sequence update)				
DT	16-OCT-2001 (Rel. 40, Last annotation update)				
DE	Negative factor (F-protein) (27 kDa protein) (3'ORF).				
GN	NEF.				
OS	Human immunodeficiency virus type 1 (BRU isolate) (HIV-1), and				
OS	Human immunodeficiency virus type 1 (New York-5 isolate) (HIV-1).				
OC	Viruses; Retroviral viruses; Retroviridae; Lentivirus.				
OX	NCBI_TaxID=11686, 11698;				
RN	[1]				
RP	SEQUENCE FROM N.A.				
RC	STRAIN-ISOLATE BRU;				
RX	MEDLINE=85099333; PubMed=2981635;				
RA	Wain-Hobson S., Sonigo P., Danos O., Cole S., Alison M.;				
RT	"Nucleotide sequence of the AIDS virus, LAV.";				
RL	Cell 40:9-17(1985).				
RN	[2]				
RP	SEQUENCE FROM N.A. (CLONE PNL4-3).				
RC	STRAIN-ISOLATE NEW YORK-5;				
RA	Buckler C.E., Buckler-White A.J., Willey R.L., McCoy J.;				
RL	Submitted (JUN-1988) to the EMBL/GenBank/DBJ databases.				
RN	[3]				
RP	POST-TRANSLATIONAL MODIFICATIONS, AND FUNCTION.				
RX	MEDLINE=88039140; PubMed=3118220;				
RA	Guy B., Kiény M.-P., Riviere Y., le Peuch C., Dott K., Girard M.;				
RT	Montagnier L., Lecocq J.-P.;				
RT	"Hiv F/3' orf encodes a phosphorylated GTP-binding protein resembling				
RT	an oncogene product.";				
RL	Nature 330:266-269(1987).				
RN	[4]				
RP	X-RAY CRYSTALLOGRAPHY (2.5 ANGSTROMS) OF 54-205 IN COMPLEX WITH NEF.				
RX	MEDLINE=96279837; PubMed=8681387;				
RA	Lee C.H., Sakela K., Mirza U.A., Chait B.T., Kuriyan J.;				
RT	"Crystal structure of the conserved core of HIV-1 Nef complexed with				
RL	a Src family SH3 domain.";				
CC	Cell 85:931-942(1996).				
CC	!- FUNCTION: NEF HAS GTPASE, GTP-BINDING AND AUTOPHOSPHORYLATING				
CC	ACTIVITIES. IT SEEMS TO DOWN-REGULATE THE CD4(T4) ANTIGEN.				
CC	-----				
CC	This SWISS-PROT entry is copyright. It is produced through a collaboration				
CC	between the Swiss Institute of Bioinformatics and the EMBL outstation -				
CC	the European Bioinformatics Institute. There are no restrictions on its				
CC	use by non-profit institutions as long as its content is in no way				
CC	modified and this statement is not removed. Usage by and for commercial				
CC	entities requires a license agreement (See http://www.isb-sib.ch/announce/				
CC	or send an email to license@isb-sib.ch).				
CC	-----				
CC	EMBL; K02013; AAB59752.1; -				
DR	EMBL; M19921; AAA44993.1; -				
DR	EMBL; A04321; CAA00353.1; -				
DR	PIR; A04008; ASLJFV				
DR	PDB; 1EFN; 11-JAN-97.				

DR HIV; K02013; NEFSBRU.  
DR HIV; M19921; NEFSNL43.  
DR InterPro; IPR001558; F-protein.  
DR Pfam; PF00469; F-protein; 1.  
KW AIDS; Myristate; GRP-binding; Phosphorylation; 3D-structure.  
FT LIPID 2 MYRISTATE.  
FT MOD\_RES 15 15 PHOSPHORYLATION (BY PKC).  
FT VARIANT 11 11 V -> I (IN CLONE PNL4-3).  
FT VARIANT 15 15 T -> A (IN CLONE PNL4-3).  
FT VARIANT 33 33 A -> V (IN CLONE PNL4-3).  
FT VARIANT 51 51 T -> N (IN CLONE PNL4-3).  
SQ SEQUENCE 206 AA; 23342 MW; 77453FC80B6004F2 CRC64;

Query Match 50.1%; Score 1130; DB 1; Length 206;  
Best Local Similarity 100.0%; Pred. No. le-86;  
Matches 206; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 110 MGGKWSKSSVVGWPTVRMRRAEPAADGVGAASRDLEKHGAITSNTAATNAACAWLEA 169  
DB 1 MGGKWSKSSVVGWPTVRMRRAEPAADGVGAASRDLEKHGAITSNTAATNAACAWLEA 60  
QY 170 QEEVEGVPVTPQVPLRPMTYKAAVDLSHFLKEKGLGSLHSQRQDILDWIYHTQGY 229  
DB 61 QEEVEGVPVTPQVPLRPMTYKAAVDLSHFLKEKGLGSLHSQRQDILDWIYHTQGY 120  
QY 230 FPDQNYTPGQVRYPLTFGWCYKLVPEPKVEEANKGENTSLHPVSLHGMDPPEV 289  
DB 121 FPDQNYTPGQVRYPLTFGWCYKLVPEPKVEEANKGENTSLHPVSLHGMDPPEV 180  
QY 290 LEWRFDSRLAFHHVARELHPEYFKNC 315  
DB 181 LEWRFDSRLAFHHVARELHPEYFKNC 206

RESULT 2  
NEF\_HV1PV  
ID NEF\_HV1PV STANDARD; PRT; 206 AA.  
AC P03405;  
DT 21-JUL-1986 (Rel. 01, Created)  
DT 21-JUL-1986 (Rel. 01, Last sequence update)  
DT 16-OCT-2001 (Rel. 40, Last annotation update)  
DE Negative factor (F-protein) (27 kDa protein) (3'ORF).  
GN NEF.  
OS Human immunodeficiency virus type 1 (PV22 isolate) (HIV-1).  
OC Viruses; Retroviral viruses; Retroviridae; Lentiviruses.  
OX NCBI\_TaxID=11700;  
RN [1]  
RP SEQUENCE FROM N.A.  
RX MEDLINE=85111157; PubMed=2982104;  
RA Muesing M.A., Smith D.H., Cabradilla C.D., Benton C.V., Lasky L.A.,  
RA Capon D.J.;  
RT "Nucleic acid structure and expression of the human  
RT AIDS/lymphadenopathy retrovirus.";  
RL Nature 313:450-458(1985).  
RN [2]  
RP POST-TRANSLATIONAL MODIFICATIONS, AND FUNCTION.  
RX MEDLINE=88039140; PubMed=3118220;  
RA Guy B., Kieny M.-P., Riviere Y., le Peuch C., Dott K., Girard M.,  
RA Montagnier L., Lecocq J.-P.;  
RT "Hiv F/3' orf encodes a phosphorylated GRP-binding protein resembling  
RT an oncogene product.";  
RL Nature 330:266-269(1987).  
CC -!- FUNCTION: NEF HAS GTPASE, GTP-BINDING AND AUTOPHOSPHORYLATING  
CC ACTIVITIES, IT SEEMS TO DOWN-REGULATE THE CD4(T4) ANTIGEN.  
CC  
CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
CC the European Bioinformatics Institute. There are no restrictions on its  
CC use by non-profit institutions as long as its content is in no way  
CC modified and this statement is not removed. Usage by and for commercial  
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).

CC EMBL; K02083; AAB59874.1; -;  
DR EMBL; X01762; -; NOT\_ANNOTATED\_CDS.  
DR PIR; A04007; ASLJVL.  
DR HSP; P03406; 1EFN.  
DR HIV; K02083; NEFSPV22.  
DR InterPro; IPR001558; F-protein.  
DR Pfam; PF00469; F-protein; 1.  
KW AIDS; Myristate; GRP-binding.  
FT LIPID 2 MYRISTATE.  
SQ SEQUENCE 206 AA; 23352 MW; EDE64281A17C6735 CRC64;

Query Match 49.4%; Score 1114; DB 1; Length 206;

Best Local Similarity 97.6%; Pred. No. 2.le-85;  
Matches 201; Conservative 3; Mismatches 2; Indels 0; Gaps 0;

QY 110 MGGKWSKSSVVGWPTVRMRRAEPAADGVGAASRDLEKHGAITSNTAATNAACAWLEA 169  
DB 1 MGGKWSKSSVVGWPAVRMRRAEPAADGVGAASRDLEKHGAITSNTAATNAACAWLEA 60  
QY 170 QEEVEGVPVTPQVPLRPMTYKAAVDLSHFLKEKGLGSLHSQRQDILDWIYHTQGY 229  
DB 61 QEEVEGVPVTPQVPLRPMTYKAAVDLSHFLKEKGLGSLHSQRQDILDWIYHTQGY 120  
QY 230 FPDQNYTPGQVRYPLTFGWCYKLVPEPKVEEANKGENTSLHPVSLHGMDPPEV 289  
DB 121 FPDQNYTPGQVRYPLTFGWCYKLVPEPKVEEANKGENTSLHPVSLHGMDPPEV 180  
QY 290 LEWRFDSRLAFHHVARELHPEYFKNC 315  
DB 181 LEWRFDSRLAFHHVARELHPEYFKNC 206

## RESULT 3

NEF\_HV1LV  
ID NEF\_HV1LV STANDARD; PRT; 206 AA.  
AC Q70627;  
DT 15-JUL-1998 (Rel. 36, Created)  
DT 15-JUL-1998 (Rel. 36, Last sequence update)  
DT 16-OCT-2001 (Rel. 40, Last annotation update)  
DE Negative factor (F-protein) (27 kDa protein) (3'ORF).  
GN NEF.  
OS Human immunodeficiency virus type 1 (WI2.3 isolate) (HIV-1).  
OC Viruses; Retroviral viruses; Retroviridae; Lentiviruses.  
OX NCBI\_TaxID=82834;  
RN [1]  
RP SEQUENCE FROM N.A.  
RX MEDLINE=95127297; PubMed=7826699;  
RA Reitz M.S. Jr., Hall L., Robert-Guroff M., Lautenberger J., Hahn B.M.,  
RA Shaw G.M., Kong L.I., Weiss S.H., Waters D., Gallo R.C., Blattner W.;  
RT "Viral variability and serum antibody response in a laboratory worker  
RT infected with HIV type 1 (HTLV type IIIB).";  
RL AIDS Res. Hum. Retroviruses 10:1143-1155(1994).  
RN [2]  
RP STRUCTURE BY NMR OF 56-206.  
RX MEDLINE=97337445; PubMed=9194185;  
RA Grzesiek S., Bax A., Hu J.S., Kaufman J., Palmer I., Stahl S.J.,  
RA Tjandra N., Wingfield P.T.;  
RT "Refined solution structure and backbone dynamics of HIV-1 Nef.";  
RL Protein Sci. 6:1248-1263(1997).  
CC -!- FUNCTION: NEF HAS GTPASE, GTP-BINDING AND AUTOPHOSPHORYLATING  
CC ACTIVITIES, IT SEEMS TO DOWN-REGULATE THE CD4(T4) ANTIGEN.  
CC  
CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
CC the European Bioinformatics Institute. There are no restrictions on its  
CC use by non-profit institutions as long as its content is in no way  
CC modified and this statement is not removed. Usage by and for commercial  
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).

```
DR PDB: 2NEF; 07-JUL-97.
DR InterPro: IPR001558; F-protein.
DR Pfam: PF00469; F-protein; 1.
KW AIDS; Myristate; GTP-binding; 3D-structure.
FT LIPID 2 MYRISTATE.
SQ SEQUENCE 206 AA; 23414 MW; ED81F68F6B61278E CRC64;

Query Match 49.3%; Score 1111; DB 1; Length 206;
Best Local Similarity 97.1%; Pred. No. 3.8e-85;
Matches 200; Conservative 5; Mismatches 1; Indels 0; Gaps 0;

QY 110 MGKWSKSVGWPTVRRMRRAEPAADGVGAASRDLEKHGAITSSNTAATNACAWLEA 169
DB 1 MGKWSKSVGWPTVRRMRRAEPAADGVGAASQDLEKHGAITSSNTAATNADCAWLEA 60
QY 170 QEEEEVGFPVTPQVPLRPMYKAAVDLSHFLKEKGGLEGLIHSQRQDILDWYHTQGY 229
DB 61 QEEEEVGFPVTPQVPLRPMYKAAVDLSHFLKEKGGLEGLIHSQRQDILDWYHTQGY 120
QY 230 FPDWQNYTPGVRYPPLTFGCWYKLVPEPKVEEANKGENTSLHPVSLHGMDPPERV 289
DB 121 FPDWQNYTPGVRYPPLTFGCWYKLVPEPKVEEANKGENTSLHPVSLHGMDPPERV 180
QY 290 LEWRFDLSLAFHHVARELHPEYFKNC 315
DB 181 LEWRFDLSLAFHHVARELHPEYFKNC 206

RESULT 4
NEF_HV112
ID NEF_HV112 STANDARD; PRT; 206 AA.
AC P04324;
DT 20-MAR-1987 (Rel. 04, Created)
DT 13-AUG-1987 (Rel. 05, Last sequence update)
DT 16-OCT-2001 (Rel. 40, Last annotation update)
DE Negative factor (F-protein) (27 kDa protein) (3'ORF).
GN NEF.
OS Human immunodeficiency virus type 1 (clone 12) (HIV-1).
OC Viruses; Retroid viruses; Retroviridae; Lentivirus.
OX NCBI_TaxID=11679;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=86177573; PubMed=3008154;
RA Arya S.K., Gallo R.C.;
RT "Three novel genes of human T-lymphotropic virus type III: immune
RT reactivity of their products with sera from acquired immune
RT deficiency syndrome patients."
RT Proc. Natl. Acad. Sci. U.S.A. 83:2209-2213(1986).
RL [2]
RN POST-TRANSLATIONAL MODIFICATIONS, AND FUNCTION.
RX MEDLINE=88039140; PubMed=3118220;
RA Guy B., Kieny M.-P., Riviere Y., le Peuch C., Dott K., Girard M.,
RA Montagnier L., Lecocq J.-P.;
RT "HIV F/3' orf encodes a phosphorylated GTP-binding protein resembling
RT an oncogene product."
RL Nature 330:266-269(1987).
CC -!- FUNCTION: NEF HAS GTPASE, GTP-BINDING AND AUTOPHOSPHORYLATING
CC ACTIVITIES, IT SEEMS TO DOWN-REGULATE THE CD4(T4) ANTIGEN.
CC
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC
CC EMBL; M11840; AAA45001.1;
DR PIR; A04006; ASLJ12.
DR HSP; P03406; LEFN.
DR HIV; M11840; NEFSPV12.
DR InterPro; IPR001558; F-protein.

DR PDB: 2NEF; 07-JUL-97.
DR InterPro: IPR001558; F-protein; 1.
DR Pfam: PF00469; F-protein; 1.
KW AIDS; Myristate; GTP-binding; 3D-structure.
FT LIPID 2 MYRISTATE.
SQ SEQUENCE 206 AA; 23414 MW; ED81F68F6B61278E CRC64;

Query Match 49.1%; Score 1108; DB 1; Length 206;
Best Local Similarity 97.1%; Pred. No. 6.7e-85;
Matches 200; Conservative 4; Mismatches 2; Indels 0; Gaps 0;

QY 110 MGKWSKSVGWPTVRRMRRAEPAADGVGAASRDLEKHGAITSSNTAATNACAWLEA 169
DB 1 MGKWSKSVGWPTVRRMRRAEPAADGVGAASRDLEKHGAITSSNTAATNACAWLEA 60
QY 170 QEEEEVGFPVTPQVPLRPMYKAAVDLSHFLKEKGGLEGLIHSQRQDILDWYHTQGY 229
DB 61 QEEEEVGFPVTPQVPLRPMYKAAVDLSHFLKEKGGLEGLIHSQRQDILDWYHTQGY 120
QY 230 FPDWQNYTPGVRYPPLTFGCWYKLVPEPKVEEANKGENTSLHPVSLHGMDPPERV 289
DB 121 FPDWQNYTPGVRYPPLTFGCWYKLVPEPKVEEANKGENTSLHPVSLHGMDPPERV 180
QY 290 LEWRFDLSLAFHHVARELHPEYFKNC 315
DB 181 LEWRFDLSLAFHHVARELHPEYFKNC 206

RESULT 5
NEF_HV1B8
ID NEF_HV1B8 STANDARD; PRT; 205 AA.
AC P05855;
DT 01-NOV-1988 (Rel. 09, Created)
DT 01-NOV-1988 (Rel. 09, Last sequence update)
DT 16-OCT-2001 (Rel. 40, Last annotation update)
DE Negative factor (F-protein) (27 kDa protein) (3'ORF).
GN NEF.
OS Human immunodeficiency virus type 1 (BH8 isolate) (HIV-1).
OC Viruses; Retroid viruses; Retroviridae; Lentivirus.
OX NCBI_TaxID=11684;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=85111123; PubMed=2578615;
RA Ratner L., Haseltine W., Patarca R., Livak K.J., Starcich B.R.,
RA Josephs S.F., Doran E.R., Rafalski J.A., Whitehorn E.A.,
RA Baumeister K., Ivanoff L., Petteway S.R. Jr., Pearson M.L.,
RA Lautenberger J.A., Papas T.S., Ghayeb J., Chang N.T., Gallo R.C.,
RA Wong-Staal F.;
RT "Complete nucleotide sequence of the AIDS virus, HTLV-III."
RL Nature 313:277-284(1985).
RN [2]
RN POST-TRANSLATIONAL MODIFICATIONS, AND FUNCTION.
RX MEDLINE=88039140; PubMed=3118220;
RA Guy B., Kieny M.-P., Riviere Y., le Peuch C., Dott K., Girard M.,
RA Montagnier L., Lecocq J.-P.;
RT "HIV F/3' orf encodes a phosphorylated GTP-binding protein resembling
RT an oncogene product."
RL Nature 330:266-269(1987).
CC -!- FUNCTION: NEF HAS GTPASE, GTP-BINDING AND AUTOPHOSPHORYLATING
CC ACTIVITIES, IT SEEMS TO DOWN-REGULATE THE CD4(T4) ANTIGEN.
CC
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC
CC EMBL; K02011; -; NOT_ANNOTATED_CDS.
DR HSP; P03406; LEFN.
DR HIV; K02011; NEFSB8.
DR InterPro; IPR001558; F-protein.
DR Pfam; PF00469; F-protein; 1.
```





```
Db 120 FPDQNTTGGVRRFLTGGWCFKLVPEKVEEANEENSGENSLHHPMSLHGMEDEPEKV 179
QY 290 LEWRFSRLAFHVAHELHPEYFKNC 315
Db 180 LVNKFDSLAFHVAHELHPEYFKNC 205

RESULT 8
NEF_HV10Y
ID NEF_HV10Y STANDARD; PRT: 211 AA.
AC P20886;
DT 01-FEB-1991 (Rel. 17, Created)
DT 01-FEB-1991 (Rel. 17, Last sequence update)
DT 16-OCT-2001 (Rel. 40, Last annotation update)
DE Negative factor (F-protein) (27 kDa protein) (3'ORF).
GN NEF.
OS Human immunodeficiency virus type 1 (HIV-1).
OC Viruses; Retroviral viruses; Retroviridae; Lentivirus.
OX NCBI_TaxID=11699;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=90148544; PubMed=2559749;
RA Huet T., Dazza M.C., Brun-Vezinet F., Roelants G.E., Wain-Hobson S.;
RT "A highly defective HIV-1 strain isolated from a healthy Gabonese
RT individual presenting an atypical western blot.";
RL AIDS 3:707-715(1989).
CC -!- FUNCTION: NEF HAS GTPASE, GTP-BINDING AND AUTOPHOSPHORYLATING
CC ACTIVITIES, IT SEEMS TO DOWN-REGULATE THE CD4(T4) ANTIGEN.
CC -!- MISCELLANEOUS: THE OYI ISOLATE WAS TAKEN FROM THE BLOOD OF A
CC HEALTHY GABONESE INDIVIDUAL.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch)
CC -----
DR EMBL: M26727; AA83398.1;
DR HSP: P03406; IEFN.
DR HIV: M26727; NEFSOYI.
DR InterPro: IPR001558; F-protein.
DR Pfam: PF00469; F-protein; 1.
KW AIDS; Myristate; GTP-binding.
FT LIPID 2 2 MYRISTATE.
SQ SEQUENCE 211 AA; 24067 MW; 64D846EEC55DAB8C CRC64;

Query Match 44.2%; Score 997.5; DB 1; Length 211;
Best Local Similarity 82.9%; Pred. No. 1e-75;
Matches 175; Conservative 22; Mismatches 9; Indels 5; Gaps 1;

QY 110 MGKWSKSSVVGWPTVRRMRRA-----EPAADGVGAASRDLEKKGAIITSSNTAATNAAC 164
Db 1 MGKWSKSKMGWPTIRERKRAELQPPPEAEGVGAASRDLEKKGAIITSSNTAATNADC 60

QY 165 AWEAQEEVEEVPVTPQVPLRPMTYKAAVDLSHFLEKKGLEGLIHSQRQDILDWY 224
Db 61 AWEAQEEVEEVPVTPQVPLRPMTYKGAALDLSHFLEKKGLEGLIYSQKQDILDWY 120

QY 225 HTQGYPPDQNTTGGVRRFLTGGWCFKLVPEKVEEANEENSGENSLHHPMSLHGMD 284
Db 121 HTQGYPPDQNTTGGVRRFLTGGWCFKLVPEKVEEANEENSGENSLHHPMSLHGMD 180

QY 285 PEREVLWRFDSRLAFHVAHELHPEYFKNC 315
Db 181 PEREVLWRFDSRLAFHVAHELHPEYFKNC 211

RESULT 9
NEF_HV1A2
```

```
ID NEF_HV1A2 STANDARD; PRT: 210 AA.
AC P03407;
DT 21-JUL-1986 (Rel. 01, Created)
DT 21-JUL-1986 (Rel. 01, Last sequence update)
DT 16-OCT-2001 (Rel. 40, Last annotation update)
DE Negative factor (F-protein) (27 kDa protein) (3'ORF).
GN NEF.
OS Human immunodeficiency virus type 1 (ARV2/SF2 isolate) (HIV-1).
OC Viruses; Retroviral viruses; Retroviridae; Lentivirus.
OX NCBI_TaxID=11685;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=85090453; PubMed=2578227;
RA Sanchez-Pescador R., Power M.D., Barr P.J., Stelmer K.S.,
RA Stenpien M.W., Brown-Shimer S.L., Gee W.W., Renard A., Randolph A.,
RA Levy J.A., Dina D., Luciw P.A.;
RT "Nucleotide sequence and expression of an AIDS-associated retrovirus
RT (ARV-2).";
RL Science 227:484-492(1985).
RN [2]
RP POST-TRANSLATIONAL MODIFICATIONS, AND FUNCTION.
RX MEDLINE=88039140; PubMed=3118220;
RA Guy B., Kieny M.-P., Riviere Y., le Peuch C., Dott K., Girard M.,
RA Montagnier L., Lecocq J.-P.;
RT "HIV F3' orf encodes a phosphorylated GTP-binding protein resembling
RT an oncogene product.";
RL Nature 330:268-269(1987).
CC -!- FUNCTION: NEF HAS GTPASE, GTP-BINDING AND AUTOPHOSPHORYLATING
CC ACTIVITIES, IT SEEMS TO DOWN-REGULATE THE CD4(T4) ANTIGEN.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch)
CC -----
DR EMBL: K02007; AA59893.1;
DR PIR: A04009; ASLJO2.
DR HSP: P03406; IEFN.
DR HIV: K02007; NEFSOYI.
DR InterPro: IPR001558; F-protein.
DR Pfam: PF00469; F-protein; 1.
KW AIDS; Myristate; GTP-binding.
FT LIPID 2 2 MYRISTATE.
SQ SEQUENCE 210 AA; 24042 MW; ED25233F8A17DAB CRC64;

Query Match 44.2%; Score 997; DB 1; Length 210;
Best Local Similarity 85.2%; Pred. No. 1e-75;
Matches 179; Conservative 16; Mismatches 11; Indels 4; Gaps 1;

QY 110 MGKWSKSSVVGWPTVRRMR-----RAEPAADGVGAASRDLEKKGAIITSSNTAATNAACA 165
Db 1 MGKWSKSKMGWGSATIRERMRRAEPAEPAADGVGAASRDLEKKGAIITSSNTAATNADCA 60

QY 166 WLEAQEEVEEVPVTPQVPLRPMTYKAAVDLSHFLEKKGLEGLIHSQRQDILDWY 225
Db 61 WLEAQEEVEEVPVTPQVPLRPMTYKAAVDLSHFLEKKGLEGLIHSQRQDILDWY 120

QY 226 TQGYPPDQNTTGGVRRFLTGGWCFKLVPEKVEEANEENSGENSLHHPMSLHGMD 285
Db 121 TQGYPPDQNTTGGVRRFLTGGWCFKLVPEKVEEANEENSGENSLHHPMSLHGMD 180

QY 286 PEREVLWRFDSRLAFHVAHELHPEYFKNC 315
Db 181 PEREVLWRFDSRLAFHVAHELHPEYFKNC 210

RESULT 10
NEF_HV1RH
ID NEF_HV1RH STANDARD; PRT: 208 AA.
```

```
AC P05858;
DT 01-NOV-1998 (Rel. 09, Created)
DT 01-NOV-1998 (Rel. 09, Last sequence update)
DT 16-OCT-2001 (Rel. 40, Last annotation update)
DE Negative factor (F-protein) (27 kDa protein) (3'ORF).
GN NEF.
OS Human immunodeficiency virus type 1 (RF/HAT isolate) (HIV-1).
OC Viruses; Retroviral viruses; Retroviridae; Lentivirus.
OX NCBI_TaxID=11701;
RN [1]
RP SEQUENCE FROM N.A.
RA Starcich B.R., Hahn B.H., Shaw G.M., McNeely P.D., Modrow S.,
RA Wolf H., Parks E.S., Parks W.P., Josephs S.F., Gallo R.C.,
RA Wong-Staal F.;
RL Submitted (XXX-1987) to the HIV data bank.
CC -!- FUNCTION: NEF HAS GTPASE, GTP-BINDING AND AUTOPHOSPHORYLATING
CC ACTIVITIES, IT SEEMS TO DOWN-REGULATE THE CD4(T4) ANTIGEN.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL; M17451; AAA45058.1; -
DR HSP; P03406; IEFN.
DR HIV; M17451; NEFSRF.
DR InterPro; IPR001558; F-protein.
DR Pfam; PF00469; F-protein; 1.
KW AIDS; Myristate; GTP-binding.
FT LIPID 2
SQ SEQUENCE 208 AA; 23532 MW; 8F836FE980F084C CRC64;
-----
Query Match 44.1%; Score 994; DB 1; Length 208;
Best Local Similarity 84.6%; Pred. No. 2e-75;
Matches 176; Conservative 16; Mismatches 14; Indels 2; Gaps 1;
QY 110 MGGKWSKSSVVGWPTVR-----EMRRAEPAADGVGAASRDLEKHGAITSNTAAATNAACAWLEA 169
DB 1 MGGKWSKSMGWPVAVREMQRAEPAADGVGAASRDLEKHGHTITSNTAANAACTWLEA 60
QY 170 Q--EEEVGFPVTPQVPLRPMTYKAAVDLSHFLKKEGGLGLIHSORRQDILDLMIYHTQ 227
DB 61 QDEDEEEVGFVPRVQVPLRPMTYKAAVDLSHFLKKEGGLGLVFSQKRDILDLWVYHTQ 120
QY 228 GFPPDMQNTPGVRYPLTFGWCYKLVPEPKVEEANKGENTSLHHPVSLHGMDDPER 287
DB 121 GFPPDMQNTPGVRYPLTFGWCYKLVPEPKVEEANKGENTSLHHPVSLHGMDDPER 180
QY 288 EVLEWRFDSRLAFHHVARELHPEYFKNC 315
DB 181 EVLVNKFDSRLAFHHVARELHPEYFKDC 208
RESULT 11
NEF_HV1Y2
ID NEF_HV1Y2 STANDARD; PRT; 214 AA.
AC P35559;
DT 01-JUN-1994 (Rel. 29, Created)
DT 01-JUN-1994 (Rel. 29, Last sequence update)
DT 16-OCT-2001 (Rel. 40, Last annotation update)
DE Negative factor (F-protein) (27 kDa protein) (3'ORF).
GN NEF.
OS Human immunodeficiency virus type 1 (YU-2 isolate) (HIV-1).
OC Viruses; Retroviral viruses; Retroviridae; Lentivirus.
OX NCBI_TaxID=36377;
RN [1]
RP SEQUENCE FROM N.A.
RA Li Y., Hui H., Burgess C.J., Price R.W., Sharp P.M., Hahn B.H.,
RX MEDLINE=93021387; PubMed=1404605;
RA
```

```
RA Shaw G.M.;
RT "Complete nucleotide sequence, genome organization, and biological
RT properties of human immunodeficiency virus type 1 in vivo: evidence
RT for limited defectiveness and complementation.";
RL J. Virol. 66:6587-6600(1992).
CC -!- FUNCTION: NEF HAS GTPASE, GTP-BINDING AND AUTOPHOSPHORYLATING
CC ACTIVITIES, IT SEEMS TO DOWN-REGULATE THE CD4(T4) ANTIGEN.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL; M93258; -; NOT_ANNOTATED_CDS.
DR PIR; I44001; I44001.
DR HSP; P03406; IEFN.
DR InterPro; IPR001558; F-protein.
DR Pfam; PF00469; F-protein; 1.
KW AIDS; Myristate; GTP-binding.
FT LIPID 2
SQ SEQUENCE 214 AA; 24532 MW; E188D43D7B084D04 CRC64;
-----
Query Match 44.0%; Score 991.5; DB 1; Length 214;
Best Local Similarity 84.2%; Pred. No. 3.3e-75;
Matches 181; Conservative 12; Mismatches 11; Indels 11; Gaps 2;
QY 110 MGGKWSKSSVVGWPTVR-----EMRRAEPAADGVGAASRDLEKHGAITSNTAA 159
DB 1 MGGKWSKSMAGWPTVRERMRRAEPAEARMRAEPAADGVGAASRDLEKHGAITSNTAA 60
QY 160 TNAACAWLEAQEEVEEVGFPVTPQVPLRPMTYKAAVDLSHFLKKEGGLGLIHSORRQDIL 219
DB 61 TNAACAWLEAQEEVEEVGFPVTPQVPLRPMTYKAAVDLSHFLKKEGGLGLIHSORRQDIL 120
QY 220 DLWYHTQGYGPDQWNYTPGPGVRYPLTFGWCYKLVPEPKVEEANKGENTSLHHPVSL 279
DB 121 DLWYHTQGYGPDQWNYTPG-GTRWPLTFGWCYKLVPEPKVEEANKGENTSLHHPVSL 179
QY 280 HGMDDPEREVLWRFDSRLAFHHVARELHPEYFKN 314
DB 180 HGMDDPEREVLWRFDSRLAFHHVARELHPEYFKN 214
RESULT 12
NEF_HV1JR
ID NEF_HV1JR STANDARD; PRT; 216 AA.
AC P20867;
DT 01-FEB-1991 (Rel. 17, Created)
DT 01-FEB-1991 (Rel. 17, Last sequence update)
DT 16-OCT-2001 (Rel. 40, Last annotation update)
DE Negative factor (F-protein) (27 kDa protein) (3'ORF).
GN NEF.
OS Human immunodeficiency virus type 1 (JRCF isolate) (HIV-1).
OC Viruses; Retroviral viruses; Retroviridae; Lentivirus.
OX NCBI_TaxID=11688;
RN [1]
RP SEQUENCE FROM N.A.
RA Koyanagi S., Chen I.S.Y.;
RL Submitted (DEC-1988) to the HIV data bank.
CC -!- FUNCTION: NEF HAS GTPASE, GTP-BINDING AND AUTOPHOSPHORYLATING
CC ACTIVITIES, IT SEEMS TO DOWN-REGULATE THE CD4(T4) ANTIGEN.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
```

```
CC EMBL; M38429; AAB03750.1; -
DR HSSP; P03406; IEFN
DR HIV; M38429; NEFSJRCFSF.
DR InterPro; IPR001558; F-protein.
DR Pfam; PF00469; F-protein; 1.
KW AIDS; Myristate; GTP-binding.
FT LIPID 2
SQ SEQUENCE 216 AA; 24567 MW; D163FFA8C71529DC CRC64;

Query Match 43.5%; Score 981; DB 1; Length 216;
Best Local Similarity 81.9%; Pred. No. 2.5e-74;
Matches 177; Conservative 15; Mismatches 14; Indels 10; Gaps 1;

Qy 110 MGGKWSKSVVGVPTVRMRRAEPAAD-----GVGAASRDLEKHGAITSNTAA 159
Dy 1 MGGKWSKSHVPGWSTVRMRRAEPAATDRVQTTEPAAGVGVASRDLEKHGAITSNTAA 60
Qy 160 TNACAWLEAOEVEEVPVTPVPLRPMYKAAVDLSHFLKEKGLGLHSQRQDIL 219
Dy 61 TNACAWLEAVEDEEVPVTPVPLRPMYKAAIDLSHFLKEKGLGLHSQRQDIL 120
Qy 220 DLWIYHTQGYFPDMQNYTPGCVRYPLTFGWCYKLVPEPKVEEANKGENTSLHVPVSL 279
Dy 121 DLWIYHTQGYFPDMQNYTAGCVRFPLTFGWCYKLVPEPKVEEANKGENTSLHVPVSL 180
Qy 280 HQMDPPEVLEWRDLSRLAHVARELHPEYFKNC 315
Dy 181 HQMDPPEVLEWRDLSRLAHVARELHPEYFKNC 216

RESULT 13
NEF_HV1S1
ID NEF_HV1S1 STANDARD; PRT; 208 AA.
AC P19546;
DT 01-FEB-1991 (Rel. 17, Created)
DT 01-FEB-1991 (Rel. 17, Last sequence update)
DT 16-OCT-2001 (Rel. 40, Last annotation update)
DE Negative factor (F-protein) (27 kDa protein) (3'ORF).
GN NEF.
OS Human immunodeficiency virus type 1 (SF162 isolate) (HIV-1).
OC Viruses; Retroviral viruses; Retroviridae; Lentivirus.
OX NCBI_TaxID=11691;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=90347835; PubMed=2384920;
RA Cheng-Mayer C., Quiroga M., Tung J.W., Dina D., Levy J.;
RT "Viral determinants of human immunodeficiency virus type 1 T-cell or
RT macrophage tropism, cytopathogenicity, and CD4 antigen modulation.";
RL J. Virol. 64:4390-4398(1990).
CC -!- FUNCTION: NEF HAS GTPASE, GTP-BINDING AND AUTOPHOSPHORYLATING
CC ACTIVITIES, IT SEEMS TO DOWN-REGULATE THE CD4(T4) ANTIGEN.
CC
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC
CC EMBL; M65024; AAA5073.1; -
DR HSSP; P03406; IEFN.
DR HIV; M38428; NEFSF162.
DR InterPro; IPR001558; F-protein.
DR Pfam; PF00469; F-protein; 1.
KW AIDS; Myristate; GTP-binding.
FT LIPID 2
SQ SEQUENCE 208 AA; 23684 MW; AUB1007D14E4E32 CRC64;

Query Match 43.2%; Score 974; DB 1; Length 208;
Best Local Similarity 79.1%; Pred. No. 6.4e-70;
Matches 163; Conservative 24; Mismatches 18; Indels 1; Gaps 1;
```

```
Best Local Similarity 84.2%; Pred. No. 9e-74;
Matches 176; Conservative 16; Mismatches 13; Indels 4; Gaps 2;

Qy 110 MGGKWSKSVVGVPTVRMRMR---AEPADGVGAASRDLEKHGAITSNTAAACAW 166
Dy 1 MGGKWSK-RMSGWSAVRMRMRRAEPAEADGVGAASRDLEKHGAITSNTAAANADCAW 59
Qy 167 LEAOEVEEVPVTPVPLRPMYKAAVDLSHFLKEKGLGLHSQRQDILDIWIHT 226
Dy 60 LEAOEVEDVGFVPRPQVPLRPMYKAAVDLSHFLKEKGLGLHSQRQDILDIWIHT 119
Qy 227 QGYFPDMQNYTPGCVRYPLTFGWCYKLVPEPKVEEANKGENTSLHVPVSLHGMDDPE 286
Dy 120 QGYFPDMQNYTPGCVRYPLTFGWCYKLVPEPKVEEANKGENTSLHVPVSLHGMDDPE 179
Qy 287 REVLEWRDLSRLAHVARELHPEYFKNC 315
Dy 180 KEVLWRPDRSLAHVARELHPEYFKNC 208

RESULT 14
NEF_HV1E1
ID NEF_HV1E1 STANDARD; PRT; 206 AA.
AC P04604;
DT 13-AUG-1987 (Rel. 05, Created)
DT 13-AUG-1987 (Rel. 05, Last sequence update)
DT 16-OCT-2001 (Rel. 40, Last annotation update)
DE Negative factor (F-protein) (27 kDa protein) (3'ORF).
GN NEF.
OS Human immunodeficiency virus type 1 (ELI isolate) (HIV-1).
OC Viruses; Retroviral viruses; Retroviridae; Lentivirus.
OX NCBI_TaxID=11689;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=86245056; PubMed=2424612;
RA Alizon M., Wain-Hobson S., Montagnier L., Sonigo P.;
RT "Genetic variability of the AIDS virus: nucleotide sequence analysis
RT of two isolates from African patients.";
RL Cell 46:63-74(1986).
RN [2]
RP POST-TRANSLATIONAL MODIFICATIONS, AND FUNCTION.
RX MEDLINE=88039140; PubMed=3118220;
RA Guy B., Kieny M.-P., Riviere Y., le Peuch C., Dott K., Girard M.,
RA Montagnier L., Lecocq J.-P.;
RT "HIV F/3' orf encodes a phosphorylated GTP-binding protein resembling
RT an oncogene product.";
RL Nature 330:266-269(1987).
CC -!- FUNCTION: NEF HAS GTPASE, GTP-BINDING AND AUTOPHOSPHORYLATING
CC ACTIVITIES, IT SEEMS TO DOWN-REGULATE THE CD4(T4) ANTIGEN.
CC
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC
CC EMBL; K03454; AAA4330.1; -
DR EMBL; A07108; CAA00617.1; -
DR HSSP; P03406; IEFN.
DR HIV; K03454; NEFSLI.
DR InterPro; IPR001558; F-protein.
DR Pfam; PF00469; F-protein; 1.
KW AIDS; Myristate; GTP-binding.
FT LIPID 2
SQ SEQUENCE 206 AA; 23612 MW; 4CFF9F18AEAB503C CRC64;

Query Match 41.1%; Score 927.5; DB 1; Length 206;
Best Local Similarity 79.1%; Pred. No. 6.4e-70;
Matches 163; Conservative 24; Mismatches 18; Indels 1; Gaps 1;
```

QY 110 MGGKWSKSSVVGWPTVRRMRRAEPAADGVGAASRDLEKKGHAITSNTAATNAACAWLEA 169  
 DB 1 MGGKWSKSSVVGWPAIRIRIRRTDPRDPAADGVGAASRDLEKKGHAITSNTAATNAACAWLEA 60  
 QY 170 QEE-EEVGFPPVTPQVPLRPMTYKAAVDLSHFLKKEGGLGSHSRRQDILDWYHTQG 228  
 DB 61 QEESEVGFPPVTPQVPLRPMTYKALDLSHFLKKEGGLGSHSRRQDILDWYHTQG 120  
 QY 229 YFPDQWNTPGGVRYPPLTFGWCYKLVPEPKVEEANKGENTSLHHPVSLHGMDDPERE 288  
 DB 121 IFPDQWNTPGGVRYPPLTFGWCYKLVPEPKVEEANKGENTSLHHPVSLHGMDDPERE 180  
 QY 289 VLEWRFDLSRLAFPHVARELHPEYFKN 314  
 DB 181 VLKWRFNLSRLAFPHVARELHPEYFKN 206

## RESULT 15

NEF\_HV126  
 ID NEF\_HV126 STANDARD; PRT; 212 AA.  
 AC P04602;  
 DT 13-AUG-1987 (Rel. 05, Last sequence update)  
 DT 13-AUG-1987 (Rel. 05, Last sequence update)  
 DT 16-OCT-2001 (Rel. 40, Last annotation update)  
 DE Negative factor (F-protein) (27 kDa protein) (3'ORF).  
 GN NEF.  
 OS Human immunodeficiency virus type 1 (Zaire 6 isolate) (HIV-1).  
 OC Viruses; Retroviral viruses; Retroviridae; Lentivirus.  
 OX NCBI\_TaxID=11708;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RX MEDLINE=87248097; PubMed=3036660;  
 RA Srinivasan A., Anand R., York D., Ranganathan P., Feorino P.,  
 RA Schochetman G., Curran J., Kalyanaraman V.S., Luciw P.A.,  
 RA Sanchez-Pescador R.;  
 RT "Molecular characterization of human immunodeficiency virus from  
 RT Zaire: nucleotide sequence analysis identifies conserved and variable  
 RT domains in the envelope gene."  
 RL Gene 52:71-82(1987).  
 RN [2]  
 RP POST-TRANSLATIONAL MODIFICATIONS, AND FUNCTION.  
 RX MEDLINE=88039140; PubMed=3118220;  
 RA Guy B., Kieny M.-P., Riviere Y., le Peuch C., Dott K., Girard M.,  
 RA Montagnier L., Lecocq J.-P.;  
 RT "HIV F3' orf encodes a phosphorylated GTP-binding protein resembling  
 RT an oncogene product."  
 RL Nature 330:266-269(1987).  
 CC -1- FUNCTION: NEF HAS GTPASE, GTP-BINDING AND AUTOPHOSPHORYLATING  
 CC ACTIVITIES, IT SEEMS TO DOWN-REGULATE THE CD4(T4) ANTIGEN.

-----  
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
 CC the European Bioinformatics Institute. There are no restrictions on its  
 CC use by non-profit institutions as long as its content is in no way  
 CC modified and this statement is not removed. Usage by and for commercial  
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announcement/>  
 CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
 CC -----

DR EMBL; K03458; AAA45381.1;  
 DR PIR; F26192; QOLJZR.  
 DR HSSP; P03406; IEFN.  
 DR HIV; K03458; NEFS26.  
 DR InterPro; IPR001558; F-protein.  
 DR Pfam; PF00469; F-protein; 1.  
 KW AIDS; Myristate; GTP-binding.  
 FT LIPID 2  
 SQ SEQUENCE 212 AA; 24415 MW; EC095A1B263047CF CRC64;

Query Match 40.4%; Score 911; DB 1; Length 212;  
 Best Local Similarity 77.4%; Pred. No. 1.6e-68;  
 Matches 164; Conservative 23; Mismatches 19; Indels 6; Gaps 2;

QY 110 MGGKWSKSSVVGWPTVRRMRRAEPAADGVGAASRDLEKKGHAITSNTAATNAAC 164  
 DB 1 MGGKWSKSSVVGWPAIRIRIRRTDPRDPAADGVGAASRDLEKKGHAITSNTAATNAAC 60  
 QY 165 AWLEAQEE-EEVGFPPVTPQVPLRPMTYKAAVDLSHFLKKEGGLGSHSRRQDILDWY 223  
 DB 61 AWLEAQEESEVGFPPVTPQVPLRPMTYKALDLSHFLKKEGGLGSHSRRQDILDWY 120  
 QY 224 YHTQGYEPDQWNTPGGVRYPPLTFGWCYKLVPEPKVEEANKGENTSLHHPVSLHGM 283  
 DB 121 YHTQGYEPDQWNTPGGVRYPPLTFGWCYKLVPEPKVEEANKGENTSLHHPVSLHGM 180  
 QY 284 DPREVLEWRFDLSRLAFPHVARELHPEYFKN 315  
 DB 181 DTEREVLEWRFDLSRLAFPHVARELHPEYFKN 212

## RESULT 16

NEF\_HV1ND  
 ID NEF\_HV1ND STANDARD; PRT; 207 AA.  
 AC P18801;  
 DT 01-NOV-1990 (Rel. 16, Created)  
 DT 01-NOV-1990 (Rel. 16, Last sequence update)  
 DT 16-OCT-2001 (Rel. 40, Last annotation update)  
 DE Negative factor (F-protein) (27 kDa protein) (3'ORF).  
 GN NEF.  
 OS Human immunodeficiency virus type 1 (NDK isolate) (HIV-1).  
 OC Viruses; Retroviral viruses; Retroviridae; Lentivirus.  
 OX NCBI\_TaxID=11695;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RX MEDLINE=90034200; PubMed=2806917;  
 RA Spire B., Sire J., Zachar V., Rey F., Barre-Sinoussi F., Galibert F.,  
 RA Hampe A., Chermann J.C.;  
 RT "Nucleotide sequence of HIV1-NDK: a highly cytopathic strain of the  
 RT human immunodeficiency virus."  
 RL Gene 81:275-284(1989).  
 CC -1- FUNCTION: NEF HAS GTPASE, GTP-BINDING AND AUTOPHOSPHORYLATING  
 CC ACTIVITIES, IT SEEMS TO DOWN-REGULATE THE CD4(T4) ANTIGEN.  
 CC -1- MISCELLANEOUS: NDK, ISOLATED FROM A ZAIRIAN PATIENT AFFECTED WITH  
 CC AIDS, AND IS A HIGHLY CYTOPATHOGENIC STRAIN.

-----  
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
 CC the European Bioinformatics Institute. There are no restrictions on its  
 CC use by non-profit institutions as long as its content is in no way  
 CC modified and this statement is not removed. Usage by and for commercial  
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announcement/>  
 CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
 CC -----

DR EMBL; M27323; AAA44874.1;  
 DR PIR; JQ0068; QOLJND.  
 DR HSSP; P03406; IEFN.  
 DR HIV; M27323; NEFSNDK.  
 DR InterPro; IPR001558; F-protein.  
 DR Pfam; PF00469; F-protein; 1.  
 KW AIDS; Myristate; GTP-binding.  
 FT LIPID 2  
 SQ SEQUENCE 207 AA; 23748 MW; 09036C2F81D45D5E CRC64;

Query Match 40.1%; Score 903.5; DB 1; Length 207;  
 Best Local Similarity 77.3%; Pred. No. 6.3e-68;  
 Matches 160; Conservative 25; Mismatches 21; Indels 1; Gaps 1;

QY 110 MGGKWSKSSVVGWPTVRRMRRAEPAADGVGAASRDLEKKGHAITSNTAATNAACAWLEA 169  
 DB 1 MGGKWSKSSVVGWPAIRIRIRRTDPAADGVGAASRDLEKKGHAITSNTAATNAACAWLEA 60  
 QY 170 QEE-EEVGFPPVTPQVPLRPMTYKAAVDLSHFLKKEGGLGSHSRRQDILDWYHTQG 228  
 DB 61 QEESEVGFPPVTPQVPLRPMTYKALDLSHFLKKEGGLGSHSRRQDILDWYHTQG 120

```
QY 229 YFPDQNTPGVRYPLTFGWCYKLVPEPKVEEANKGENTSLHPVSLHGMDDPERE 288
DB 121 IFPDQNTPGGIRYPLTFGWCYKLVPEPKVEEANKGENTSLHPVSLHGMDDPERE 180
QY 289 VLEWRDPSRLAFHVAHRELHPEYFKN 315
DB 181 VLMWRNSRLALEKHARELHPEYFKDC 207

RESULT 17
NEF_HV104
ID NEF_HV104 STANDARD; PRT; 205 AA.
AC P24741;
DT 01-MAR-1992 (Rel. 21, Created)
DT 01-MAR-1992 (Rel. 21, Last sequence update)
DT 16-OCT-2001 (Rel. 40, Last annotation update)
DE Negative factor (F-protein) (27 kDa protein) (3'ORF).
NEF.
GN NEF.
OS Human immunodeficiency virus type 1 (strain Ugandan / isolate U455)
OC (HIV-1).
OC Viruses; Retroviridae; Lentivirus.
OX NCBI_TaxID=11703;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=9109081; PubMed=2265025;
RA Oram J.D., Downing R.G., Roff M., Clegg J.C.S., Serwadda D.,
RA Carswell J.W.;
RT "Nucleotide sequence of a Ugandan HIV-1 provirus reveals genetic
RT diversity from other HIV-1 isolates."
RL AIDS Res. Hum. Retroviruses 6:1073-1078(1990).
CC -1- FUNCTION: NEF HAS GTPASE, GTP-BINDING AND AUTOPHOSPHORYLATING
CC ACTIVITIES, IT SEEMS TO DOWN-REGULATE THE CD4(T4) ANTIGEN.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (see http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
CC EMBL; M62320; AAA75023.1; -
CC HSP; P03406; IEFN.
CC InterPro: IPR001558; F-protein.
CC Pfam: PF00469; F-protein; 1.
CC AIDS; Myristate; Gtp-binding.
CC LIPID 2
CC MYRISTATE (BY SIMILARITY).
CC SEQUENCE 205 AA; 23253 MW; 869AB03B6E7893C4 CRC64;

Query Match 38.7%; Score 872; DB 1; Length 205;
Best Local Similarity 75.1%; Pred. No. 2.6e-65;
Matches 154; Conservative 24; Mismatches 27; Indels 0; Gaps 0;

QY 110 MGGKWSKSSVVGWPTVRERMRRAEPAADGVGAASRDLEKHGAITSNTAATNAACAWLEA 169
DB 1 MGGKWSKSRVWEPVRKMRMTAAAGVGAVSODLDKYGAVTSSNTSSNASCWALEA 60
QY 170 QEEVEVGFPVTPQVPLRPMTYKAAVDLSHFLKEKGLGLHSQRQDILDWYHTQGY 229
DB 61 QEEGDVGFPVTPQVPLRPMTYKAAFDLSFFLKEKGLGLHSQRQDILDWYHTQGF 120
QY 230 FPDQNTPGVRYPLTFGWCYKLVPEPKVEEANKGENTSLHPVSLHGMDDPERE 289
DB 121 FPDQNTPGGIRYPLTFGWCYKLVPEPKVEEANKGENTSLHPVSLHGMDDPERE 180
QY 290 LEWRDPSRLAFHVAHRELHPEYFKN 314
DB 181 LMKFPDSTALKHAYELHPEYFKD 205

RESULT 18
NEF_HV1MN
ID NEF_HV1MN STANDARD; PRT; 205 AA.
AC P05856;
DT 01-NOV-1988 (Rel. 09, Created)
DT 01-NOV-1988 (Rel. 09, Last sequence update)
DT 16-OCT-2001 (Rel. 40, Last annotation update)
DE Negative factor (F-protein) (27 kDa protein) (3'ORF).
NEF.
GN NEF.
OS Human immunodeficiency virus type 1 (MN isolate) (HIV-1).
OC Viruses; Retroviridae; Lentivirus.
OX NCBI_TaxID=11696;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=88219542; PubMed=3369091;
RA Gurgo C., Guo H.-G., Franchini G., Aldovini A., Colliatti E.,
RA Farrell K., Wong-Staal F., Gallo R.C., Reitz M.S. Jr.;
RT "Envelope sequences of two new United States HIV-1 isolates."
RL Virology 164:531-536(1988).
CC -1- FUNCTION: NEF HAS GTPASE, GTP-BINDING AND AUTOPHOSPHORYLATING
CC ACTIVITIES, IT SEEMS TO DOWN-REGULATE THE CD4(T4) ANTIGEN.
CC -1- MISCELLANEOUS: THE MN ISOLATE WAS TAKEN FROM A PEDIATRIC AIDS
CC PATIENT IN 1984.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (see http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
CC EMBL; M17449; AAA44858.1; -
CC HSP; P03406; IEFN.
CC HIV; M17449; NEFSMN.
CC InterPro: IPR001558; F-protein.
CC Pfam: PF00469; F-protein; 1.
CC AIDS; Myristate; Gtp-binding.
CC LIPID 2
CC MYRISTATE (BY SIMILARITY).
CC SEQUENCE 182 AA; 20632 MW; 2C5B5A6E3EF13B26 CRC64;

Query Match 37.3%; Score 842; DB 1; Length 182;
Best Local Similarity 87.4%; Pred. No. 6.8e-63;
Matches 153; Conservative 11; Mismatches 7; Indels 4; Gaps 2;

QY 110 MGGKWSKSSVVGWPTVRERMRRAEPAADGVGAASRDLEKHGAITSNTAATNAACAW 166
DB 1 MGGKWSK-RVTPGWPVTRERMRRAEPAELADGVGAASRDLEKHGAITSNTAATNAACAW 59
QY 167 LEAQEEVEVGFPVTPQVPLRPMTYKAAVDLSHFLKEKGLGLHSQRQDILDWYHT 226
DB 60 LEAQEEVEVGFPVTPQVPLRPMTYKAAVDLSHFLKEKGLGLHSQRQDILDWYHT 119
QY 227 QGYFPDQNTPGVRYPLTFGWCYKLVPEPKVEEANKGENTSLHPVSLHGMDDPERE 281
DB 120 QGYFPDQNTPGGIRYPLTFGWCYKLVPEPKVEEANKGENTSLHPVSLHGMDDPERE 174

RESULT 19
NEF_HV12H
ID NEF_HV12H STANDARD; PRT; 205 AA.
AC P05859;
DT 01-NOV-1988 (Rel. 09, Created)
DT 01-NOV-1988 (Rel. 09, Last sequence update)
DT 16-OCT-2001 (Rel. 40, Last annotation update)
DE Negative factor (F-protein) (27 kDa protein) (3'ORF).
NEF.
GN NEF.
OS Human immunodeficiency virus type 1 (Zaire H2321 isolate) (HIV-1).
OC Viruses; Retroviridae; Lentivirus.
OX NCBI_TaxID=11692;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=89228766; PubMed=2713163;
RA Srinivasan A., York D., Butler D. Jr., Jannoun-Nasr R., Getchell J.,
```

```
RA McCormick J., Ou C.Y., Myers G., Smith T., Chen E.;  
RT "Molecular characterization of HIV-1 isolated from a serum collected  
RT in 1976: nucleotide sequence comparison to recent isolates and  
RT generation of hybrid HIV-1";  
RL AIDS Res. Hum. Retroviruses 5:121-129(1989).  
CC -1- FUNCTION: NEF HAS GTPASE, GTP-BINDING AND AUTOPHOSPHORYLATING  
CC ACTIVITIES, IT SEEMS TO DOWN-REGULATE THE CD4(T4) ANTIGEN.  
CC -----  
CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
CC the European Bioinformatics Institute. There are no restrictions on its  
CC use by non-profit institutions as long as its content is in no way  
CC modified and this statement is not removed. Usage by and for commercial  
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/  
CC or send an email to license@isb-sib.ch).  
CC -----  
DR EMBL: M15896; AAB53951.1; -;  
DR PIR: B44963;  
DR HSP: P03406; IEFN.  
DR HIV: M15896; NEFS2321.  
DR InterPro: IPR001558; F-protein.  
DR Pfam: PF00469; F-protein; 1.  
DR AIDS; Myristate; GTP-binding.  
KW AIDS; Myristate; GTP-binding.  
FT LIPID 2  
SQ SEQUENCE 205 AA; 23306 MW; EA55B18AF412A9D0 CRC64;  
-----  
Query Match 37.1%; Score 837; DB 1; Length 205;  
Best Local Similarity 73.2%; Pred. No. 2.1e-62;  
Matches 153; Conservative 23; Mismatches 25; Indels 8; Gaps 3;  
QY 110 MGGKWSKSSVVGWTVTRMRMR---AEPADGVGAASRDLEKKGAISSNTAATNAACAW 166  
DB 1 MGNKWSK----GWPAVRERIRQTTPAPPAEGVGAASQDLAKHGAISSNTATNPPDCAW 56  
QY 167 LEAQEESEVGGFPVTPGVRVPLTFGCWYKLVPEPDKVEEANKGENTSLLLHPVSLHGMDDP 225  
DB 57 LEAQEESEVGGFPVTPGVRVPLTFGCWYKLVPEPDKVEEANKGENTSLLLHPVSLHGMDDP 116  
QY 226 TQGFPPDQNTTPGVRVPLTFGCWYKLVPEPDKVEEANKGENTSLLLHPVSLHGMDDP 285  
DB 117 TQGFPPDQNTTPGVRVPLTFGCWYKLVPEPDKVEEANKGENTSLLLHPVSLHGMDDP 176  
QY 286 EREVLEWRFDSSLAPHHVARELHPEYKNC 314  
DB 177 EREVLEWRFDSSLAPHHVARELHPEYKNC 205  
-----  
RESULT 20  
NEF_HV1MA STANDARD: PRT; 209 AA.  
ID NEF_HV1MA  
AC P04603;  
DT 13-AUG-1987 (Rel. 05, Created)  
DT 13-AUG-1987 (Rel. 05, Last sequence update)  
DT 16-OCT-2001 (Rel. 40, Last annotation update)  
DE Negative factor (F-protein) (27 kDa protein) (3'ORF).  
GN NEF.  
OS Human immunodeficiency virus type 1 (MAL isolate) (HIV-1).  
OC Viruses; Retroviral viruses; Retroviridae; Lentivirus.  
OX NCBI_TaxID=11697;  
RN [1]  
RP SEQUENCE FROM N.A.  
RX MEDLINE=86245056; PubMed=2424612;  
RA Allizon M., Wain-Hobson S., Montagnier L., Sonigo P.;  
RT "Genetic variability of the AIDS virus: nucleotide sequence analysis  
RT of two isolates from African patients.";  
RL Cell 46:63-74(1986).  
RN [2]  
RP POST-TRANSLATIONAL MODIFICATIONS, AND FUNCTION.  
RX MEDLINE=88039140; PubMed=3118220;  
RA Guy B., Kieny M.-P., Riviere Y., le Peuch C., Dott K., Girard M.,  
RA Montagnier L., Lecocq J.-P.;  
RT "HIV F3' orf encodes a phosphorylated GTP-binding protein resembling  
RT an oncogene product.";  
RL Nature 330:266-269(1987).  
CC -1- FUNCTION: NEF HAS GTPASE, GTP-BINDING AND AUTOPHOSPHORYLATING  
CC ACTIVITIES, IT SEEMS TO DOWN-REGULATE THE CD4(T4) ANTIGEN.  
CC -----  
CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
CC the European Bioinformatics Institute. There are no restrictions on its  
CC use by non-profit institutions as long as its content is in no way  
CC modified and this statement is not removed. Usage by and for commercial  
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/  
CC or send an email to license@isb-sib.ch).  
CC -----  
DR EMBL: X04415; CAA28017.1; -;  
DR EMBL: A07116; CAA00624.1; -;  
DR HSP: P03406; IEFN.  
DR HIV: K03456; NEFSMAL.  
DR InterPro: IPR001558; F-protein.  
DR Pfam: PF00469; F-protein; 1.  
DR AIDS; Myristate; GTP-binding.  
KW AIDS; Myristate; GTP-binding.  
FT LIPID 2  
SQ SEQUENCE 209 AA; 23644 MW; DOB30A2442C8CC44 CRC64;  
-----  
Query Match 37.0%; Score 834.5; DB 1; Length 209;  
Best Local Similarity 70.1%; Pred. No. 3.4e-62;  
Matches 148; Conservative 26; Mismatches 30; Indels 7; Gaps 2;  
QY 110 MGGKWSKSSVVGWTVTRMRMRRAEPADGVG-----AASRDLEKKGAISSNTAATNAAC 164  
DB 1 MGGKWSKSSVVGWTVTRMRMRRAEPADGVG-----AASRDLEKKGAISSNTAATNAAC 60  
QY 165 AWLEAQEESEVGGFPVTPGVRVPLTFGCWYKLVPEPDKVEEANKGENTSLLLHPVSLHGMDD 224  
DB 61 E--PPEEESEVGGFPVTPGVRVPLTFGCWYKLVPEPDKVEEANKGENTSLLLHPVSLHGMDD 118  
QY 225 HTQGFPPDQNTTPGVRVPLTFGCWYKLVPEPDKVEEANKGENTSLLLHPVSLHGMDD 284  
DB 119 HTQGFPPDQNTTPGVRVPLTFGCWYKLVPEPDKVEEANKGENTSLLLHPVSLHGMDD 178  
QY 285 PEREVLWRFDSSLAPHHVARELHPEYKNC 315  
DB 179 AEREVLWRFDSSLAPHHVARELHPEYKNC 209  
-----  
RESULT 21  
NEF_SIVCZ STANDARD: PRT; 205 AA.  
ID NEF_SIVCZ  
AC P17664;  
DT 01-AUG-1990 (Rel. 15, Created)  
DT 01-AUG-1990 (Rel. 15, Last sequence update)  
DT 16-OCT-2001 (Rel. 40, Last annotation update)  
DE Negative factor (F-protein) (27 kDa protein) (3'ORF).  
GN NEF.  
OS Chimpanzee immunodeficiency virus (SIV(cpz)) (CIV).  
OC Viruses; Retroviral viruses; Retroviridae; Lentivirus.  
OX NCBI_TaxID=11723;  
RN [1]  
RP SEQUENCE FROM N.A.  
RX MEDLINE=90259077; PubMed=2188136;  
RA Huet T., Cheynier R., Meyerhans A., Roelants G., Wain-Hobson S.;  
RT "Genetic organization of a chimpanzee lentivirus related to HIV-1.";  
RL Nature 345:356-359(1990).  
CC -1- FUNCTION: NEF HAS GTPASE, GTP-BINDING AND AUTOPHOSPHORYLATING  
CC ACTIVITIES, IT SEEMS TO DOWN-REGULATE THE CD4(T4) ANTIGEN.  
CC -----  
CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
CC the European Bioinformatics Institute. There are no restrictions on its  
CC use by non-profit institutions as long as its content is in no way  
CC modified and this statement is not removed. Usage by and for commercial  
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/  
CC or send an email to license@isb-sib.ch).
```



```
CC EMBL: X52154; CAA36408.1; -
DR PIR: S09991; ASLJIK.
DR HSP: P03406; IEFN.
DR HIV: X52154; NEFSCP2.
DR InterPro: IPR001558; F-protein.
DR Pfam: PF00469; F-protein; 1.
DR KW: Myristate; GTP-binding.
DR AIDS: Myristate; GTP-binding.
DR LIPID 2
FT SEQUENCE 205 AA; 23850 MW; 21E0A3EC99F1811F CRC64;
SQ

Query Match 36.9%; Score 832; DB 1; Length 205;
Best Local Similarity 70.28; Pred. No. 5.3e-62;
Matches 144; Conservative 27; Mismatches 34; Indels 0; Gaps 0;

QY 110 MGGKWSKSSVVGWPTVRRMRRAEPAADGVGAASRDLEKKGGLHSHFSLHSGMDDPER 169
D 1 MGTKWSKSSLVGWPEVRRRIEAPTAAGVGVEVSKOLERHCAITSRNPTETNQTALWLEE 60
QY 170 QEEVEVGFVPTQVPLRPMYTKAAVDLSHFLKKEGGLHSHFSLHSGMDDPER 229
D 61 MDNEEVGFVPTQVPLRPMYTKAAVDLSHFLKKEGGLHSHFSLHSGMDDPER 120
QY 230 FPDQNTYTPGCVRYPLTFEGWCYKLVPEVDPKVEEANKGENTSLHHPVSLHGMDDPER 289
D 121 FPDQNTYTPGCVRYPLTFEGWCYKLVPEVDPKVEEANKGENTSLHHPVSLHGMDDPER 180
QY 290 LEWRFDRLAFHVAHRELHPEYFKN 314
D 181 LEWRFDRLAFHVAHRELHPEYFKN 205

RESULT 22
NEF_HV1SC STANDARD; PRT; 239 AA.
AC P05857;
DT 01-NOV-1988 (Rel. 09, Created)
DT 01-NOV-1988 (Rel. 09, Last sequence update)
DT 16-OCT-2001 (Rel. 40, Last annotation update)
DE Negative factor (F-protein) (27 kDa protein) (3'ORF).
GN NEF.
OS Human immunodeficiency virus type 1 (SC isolate) (HIV-1).
OC Viruses; Retroviral viruses; Retroviridae; Lentivirus.
OX NCBI_TaxID=11702;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=88219542; PubMed=3369091;
RA Gurgu C., Guo H.-G., Franchini G., Aldovini A., Collalti E.,
RA Farrell K., Wong-Staal F., Gallo R.C., Reitz M.S. Jr.;
RT "Envelope sequences of two new United States HIV-1 isolates.";
RL Virology 164:531-536(1988).
CC -1- FUNCTION: NEF HAS GTPASE, GTP-BINDING AND AUTOPHOSPHORYLATING
CC ACTIVITIES, IT SEEMS TO DOWN-REGULATE THE CD4(T4) ANTIGEN.
CC -1- MISCELLANEOUS: THE SC ISOLATE WAS TAKEN FROM AN ARC PATIENT IN
CC 1984 IN SOUTHERN CALIFORNIA.
CC
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (see http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC
CC EMBL: M17450; AAA45064.1; -
DR HSP: P03406; IEFN.
DR HIV: M17450; NEFSSC.
DR InterPro: IPR001558; F-protein.
DR Pfam: PF00469; F-protein; 1.
DR KW: Myristate; GTP-binding.
DR LIPID 2
FT SEQUENCE 239 AA; 26799 MW; 704A17E54763A99B CRC64;
SQ
```

```
Query Match 36.3%; Score 819; DB 1; Length 239;
Best Local Similarity 75.5%; Pred. No. 7.8e-61;
Matches 157; Conservative 16; Mismatches 31; Indels 4; Gaps 2;

QY 110 MGGKWSKSSVVGWPTVRRMRRAEPAADGVGAASRDLEKKGGLHSHFSLHSGMDDPER 169
D 1 MGGKWSKSSVVGWPTVRRMRRAEPAADGVGAASRDLEKKGGLHSHFSLHSGMDDPER 60
QY 170 QEEVEVGFVPTQVPLRPMYTKAAVDLSHFLKKEGGLHSHFSLHSGMDDPER 229
D 61 QEEVEVGFVPTQVPLRPMYTKAAVDLSHFLKKEGGLHSHFSLHSGMDDPER 118
QY 230 FPDQNTYTPGCVRYPLTFEGWCYKLVPEVDPKVEEANKGENTSLHHPVSLHGMDDPER 287
D 119 ATSLIGRTHGQGSIDPLCFGWCYKLVPEVDPKVEEANKGENTSLHHPVSLHGMDDPER 178
QY 288 EVLEWRFDRLAFHVAHRELHPEYFKN 315
D 179 EVLEWRFDRLAFHVAHRELHPEYFKN 206

RESULT 23
NEF_HV1H2 STANDARD; PRT; 123 AA.
AC P04601; O09780;
DT 13-AUG-1987 (Rel. 05, Created)
DT 15-JUL-1999 (Rel. 38, Last sequence update)
DT 16-OCT-2001 (Rel. 40, Last annotation update)
DE Negative factor (F-protein) (27 kDa protein) (3'ORF).
GN NEF.
OS Human immunodeficiency virus type 1 (HXB2 isolate) (HIV-1).
OC Viruses; Retroviral viruses; Retroviridae; Lentivirus.
OX NCBI_TaxID=11706;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=87299196; PubMed=3040055;
RA Ratner L., Fisher A., Jagodzinski L.L., Mitsuya H., Liou R.-S.,
RA Gallo R.C., Wong-Staal F.;
RT "Complete nucleotide sequences of functional clones of the AIDS
RT virus.";
RL AIDS Res. Hum. Retroviruses 3:57-69(1987).
RN [2]
RP REVISIONS.
RA Ratner L., Fisher A., Jagodzinski L.L., Mitsuya H., Liou R.-S.,
RA Gallo R.C., Wong-Staal F.;
RL Submitted (APR-1997) to the EMBL/GenBank/DBJ databases.
RN [3]
RP POST-TRANSLATIONAL MODIFICATIONS, AND FUNCTION.
RX MEDLINE=88039140; PubMed=3118220;
RA Guy B., Kieny M.-P., Riviere Y., le Peuch C., Dott K., Girard M.,
RA Montagnier L., Lecocq J.-P.;
RT "HIV F/3' orf encodes a phosphorylated GTP-binding protein resembling
RT an oncogene product.";
RL Nature 330:266-269(1987).
CC -1- FUNCTION: NEF HAS GTPASE, GTP-BINDING AND AUTOPHOSPHORYLATING
CC ACTIVITIES, IT SEEMS TO DOWN-REGULATE THE CD4(T4) ANTIGEN.
CC -1- MISCELLANEOUS: IN THIS ISOLATE A MUTATION IN POSITION 124 ADDS A
CC STOP CODON AND THE PROTEIN IS MUCH SHORTER THAN IN OTHER ISOLATES
CC (210 AA).
CC
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (see http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC
CC EMBL: K03455; AAB50263.1; -
DR EMBL: AF033819; AAC82597.1; -
DR HSP: Q70627; 2NEF.
```





Infect. Immun. 61:4546-4552(1993).

[4]

SEQUENCE FROM N.A.

STRAIN=EAGAN, 3639, 3640, NCTC 8468, 6-7636, AND HK695;

MEDLINE=95122210; PubMed=7822043;

Song X.-M., Forsgren A., Janson H.;

"The gene encoding protein B (hpd) is highly conserved among

Haemophilus influenzae type B and nontypeable strains.";

Infect. Immun. 63:696-699(1995).

[5]

CHARACTERIZATION.

STRAIN=NTHI 772;

MEDLINE=92192801; PubMed=1548059;

Janson H., Heden L.O., Forsgren A.;

"Protein B, the immunoglobulin D-binding protein of Haemophilus

influenzae, is a lipoprotein.";

Infect. Immun. 60:1336-1342(1992).

-I- FUNCTION: GLYCEROPHOSPHORYL DIESTER PHOSPHODIESTERASE HYDROLYZES

DEACETYLATED PHOSPHOLIPIDS TO G3P AND THE CORRESPONDING ALCOHOLS.

HAS A SPECIFIC AFFINITY FOR HUMAN IMMUNOGLOBULIN D MYELOMA

PROTEIN.

-I- CATALYTIC ACTIVITY: A glycerophosphodiester + H(2)O = an alcohol +

sn-glycerol 3-phosphate.

-I- SUBCELLULAR LOCATION: Attached to the outer membrane by a lipid

anchor.

-I- PTM: CONTAINS BOTH ESTER- AND AMIDE-LINKED FATTY ACIDS.

-I- MISCELLANEOUS: THE SEQUENCE SHOWN IS THAT OF STRAINS NTHI 772 AND

RD / KW20.

-I- SIMILARITY: TO E.COLI AND B.SUBTILIS GLPO.

-----

This SWISS-PROT entry is copyright. It is produced through a collaboration

between the Swiss Institute of Bioinformatics and the EMBL outstation -

the European Bioinformatics Institute. There are no restrictions on its

use by non-profit institutions as long as its content is in no way

modified and this statement is not removed. Usage by and for commercial

entities requires a license agreement (See <http://www.isb-sib.ch/announce/>

or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).

-----

EMBL: U32751; AAC22348.1; -

EMBL: M37487; AAA24998.1; -

EMBL: L12445; AAA24999.1; -

EMBL: Z35656; CAA84715.1; -

EMBL: Z35657; CAA84716.1; -

EMBL: Z35658; CAA84717.1; -

EMBL: Z35659; CAA84718.1; -

EMBL: Z35660; CAA84719.1; -

EMBL: Z35661; CAA84720.1; -

TIGR: H10689; -

InterPro: IPR0041129; GDFD.

Pfam: PF03009; GDFP; 1.

DR DR PROSITE; PS00013; PROKAR\_LIPOPROTEIN; 1.

KW GlyceroL metabolism; Hydrolase; Signal; Lipoprotein; Outer membrane;

Complete proteome.

FT SIGNAL 1 18

FT CHAIN 19 364

FT

FT LIPID 19 19 GLYCEROPHOSPHORYL DIESTER

FT PHOSPHODIESTERASE.

FT N-ACYL DIGLYCERIDE.

FT A -> T (IN STRAIN NCTC 8468).

FT L -> V (IN STRAIN NCTC 8468).

FT N -> S (IN STRAIN NCTC 8468).

FT N -> K (IN STRAIN 6-7626).

FT D -> H (IN STRAIN NCTC 8468).

FT H -> Q (IN STRAINS EAGAN; 3639; 3640;

FT 6-7626; HK695; MINNA)

FT S -> A (IN STRAINS EAGAN; 3639; 3640;

FT NCTC 8468; 6-7626; HK695; MINNA).

FT Y -> H (IN STRAINS EAGAN; 3639; 3640;

FT NCTC 8468; 6-7626; HK695; MINNA).

FT R -> H (IN STRAIN NCTC 8468).

FT K -> Q (IN STRAIN 6-7626).

FT K -> R (IN STRAIN 6-7626).

FT T -> A (IN STRAINS EAGAN; 3639; 3640;

FT NCTC 8468; 6-7626; HK695; MINNA).

FT

```
CC TRANS-ACTIVATING RESPONSIVE SEQUENCE (TAR) RNA ELEMENT AND
CC ACTIVATES TRANSCRIPTION INITIATION AND/OR ELONGATION FROM THE LTR
CC PROMOTER.
CC -1- SUBUNIT: BINDS CYCLIN T1 (BY SIMILARITY).
CC -1- SUBCELLULAR LOCATION: NUCLEAR; NUCLEOLAR.
CC -1- MISCELLANEOUS: THE TWO ISOLATES TAT SEQUENCE ARE IDENTICAL.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation.
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
CC EMBL: M14100; AAA44676.1; -.
CC EMBL: M15654; TAT5BH102.
CC HIV: M14100; TAT5HXB3.
CC InterPro: IPR001831; HIV_Tat.
CC Pfam: PF00539; Tat; 1.
CC PRINTS: PR00055; HIVTATDOMAIN.
CC Transcription regulation; Activator; RNA-binding; Nuclear protein;
CC AIDS.
CC SEQUENCE 86 AA; 9784 MW; 4DD609415FAF9015 CRC64;

Query Match 21.6%; Score 486; DB 1; Length 86;
Best Local Similarity 100.0%; Pred. No. 8.9e-34;
Matches 85; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 318 EPVDPRLPWPWKGSGQPKTACTNCYKCKCCFHCQVCFTKALGISYGRKKRRRPPQG 377
DB 2 EPVDPRLPWPWKGSGQPKTACTNCYKCKCCFHCQVCFTKALGISYGRKKRRRPPQG 61

QY 378 SOTHQVSLSKQPTSGRSGDPTGPK 402
DB 62 SOTHQVSLSKQPTSGRSGDPTGPK 86

RESULT 27
TAT_HVIPV
ID TAT_HVIPV STANDARD; PRT; 86 AA.
AC P04607;
DT 13-AUG-1987 (Rel. 05, Created)
DT 13-AUG-1987 (Rel. 05, Last sequence update)
DT 16-OCT-2001 (Rel. 40, Last annotation update)
DE TAT protein (Transactivating regulatory protein).
GN TAT.
OS Human immunodeficiency virus type 1 (PV22 isolate) (HIV-1).
OC Viruses; Retroid viruses; Retroviridae; Lentivirus.
OX NCBI_TaxID=11700;
RN [1]
RP SEQUENCE FROM N.A.
RA MEDLINE=85111157; PubMed=2982104;
RX Muesing M.A., Smith D.H., Cabradilla C.D., Benton C.V., Lasky L.A.,
RA Capon D.J.
RT "Nucleic acid structure and expression of the human
RT Aids/lymphadenopathy retrovirus."
RL Nature 313:450-458(1985).
CC -1- FUNCTION: TRANSCRIPTIONAL REGULATOR THAT ACTS BY BINDING TO THE
CC TRANS-ACTIVATING RESPONSIVE SEQUENCE (TAR) RNA ELEMENT AND
CC ACTIVATES TRANSCRIPTION INITIATION AND/OR ELONGATION FROM THE LTR
CC PROMOTER.
CC -1- SUBUNIT: BINDS CYCLIN T1 (BY SIMILARITY).
CC -1- SUBCELLULAR LOCATION: NUCLEAR; NUCLEOLAR.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation.
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
```

```
CC or send an email to license@isb-sib.ch).
CC -----
CC EMBL: K02083; AAB59870.1; -.
CC EMBL: X01762; -. NOT_ANNOTATED_CDS.
CC HIV: K02083; TATSPV22.
CC InterPro: IPR001831; HIV_Tat.
CC Pfam: PF00539; Tat; 1.
CC PRINTS: PR00055; HIVTATDOMAIN.
CC Transcription regulation; Activator; RNA-binding; Nuclear protein;
CC AIDS.
CC SEQUENCE 86 AA; 9794 MW; 4DD3C6415FAF9015 CRC64;

Query Match 21.3%; Score 481; DB 1; Length 86;
Best Local Similarity 98.8%; Pred. No. 2.3e-33;
Matches 84; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 318 EPVDPRLPWPWKGSGQPKTACTNCYKCKCCFHCQVCFTKALGISYGRKKRRRPPQG 377
DB 2 EPVDPRLPWPWKGSGQPKTACTNCYKCKCCFHCQVCFTKALGISYGRKKRRRPPQG 61

QY 378 SOTHQVSLSKQPTSGRSGDPTGPK 402
DB 62 SOTHQVSLSKQPTSGRSGDPTGPK 86

RESULT 28
TAT_HV112
ID TAT_HV112 STANDARD; PRT; 86 AA.
AC P04326;
DT 20-MAR-1987 (Rel. 04, Created)
DT 13-AUG-1987 (Rel. 05, Last sequence update)
DT 16-OCT-2001 (Rel. 40, Last annotation update)
DE TAT protein (Transactivating regulatory protein).
GN TAT.
OS Human immunodeficiency virus type 1 (clone 12) (HIV-1).
OC Viruses; Retroid viruses; Retroviridae; Lentivirus.
OX NCBI_TaxID=11679;
RN [1]
RP SEQUENCE FROM N.A.
RA MEDLINE=86177573; PubMed=3008154;
RX Arya S.K., Gallo R.C.;
RT "Three novel genes of human T-lymphotropic virus type III: Immune
RT reactivity of their products with sera from acquired immune
RT deficiency syndrome patients."
RL Proc. Natl. Acad. Sci. U.S.A. 83:2209-2213(1986).
CC -1- FUNCTION: TRANSCRIPTIONAL REGULATOR THAT ACTS BY BINDING TO THE
CC TRANS-ACTIVATING RESPONSIVE SEQUENCE (TAR) RNA ELEMENT AND
CC ACTIVATES TRANSCRIPTION INITIATION AND/OR ELONGATION FROM THE LTR
CC PROMOTER.
CC -1- SUBUNIT: BINDS CYCLIN T1 (BY SIMILARITY).
CC -1- SUBCELLULAR LOCATION: NUCLEAR; NUCLEOLAR.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation.
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
CC EMBL: M11840; AAA44999.1; -.
CC HIV: M11840; TATSPCV12.
CC InterPro: IPR001831; HIV_Tat.
CC Pfam: PF00539; Tat; 1.
CC PRINTS: PR00055; HIVTATDOMAIN.
CC Transcription regulation; Activator; RNA-binding; Nuclear protein;
CC AIDS.
CC SEQUENCE 86 AA; 9758 MW; 4DD609414FBE9115 CRC64;

Query Match 21.2%; Score 478; DB 1; Length 86;
```

Best Local Similarity 98.8%; Pred. No. 4.1e-33;  
Matches 84; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 318 EPVDPRLPEWKHPGSPKTCACNYCKCCFHCQVCFITKALGISYGRKKRRRRPPQG 377  
|||||  
Db 2 EPVDPRLPEWKHPGSPKTCACNYCKCCFHCQVCFITKALGISYGRKKRRRRPPQG 61  
|||||  
QY 378 SOTHQVSLSKQPTSQSRGDPGPK 402  
|||||  
Db 62 SOTHQVSLSKQPTSQSRGDPGPK 86  
|||||

## RESULT 29

TAT\_HV1BR STANDARD: PRT: 86 AA.  
ID TAT\_HV1BR  
AC P04610;  
DT 13-AUG-1987 (Rel. 05, Created)  
DT 13-AUG-1987 (Rel. 05, Last sequence update)  
DT 16-OCT-2001 (Rel. 40, Last annotation update)  
DE TAT protein (Transactivating regulatory protein).  
GN TAT.  
OS Human immunodeficiency virus type 1 (BRU isolate) (HIV-1).  
OC Viruses; Retroviral viruses; Retroviridae; Lentivirus.  
OX NCBI\_TaxID=11686;  
RN [1]  
RP SEQUENCE FROM N.A.  
RX MEDLINE=85099333; PubMed=2981635;  
RA Wain-Hobson S., Sonigo P., Danos O., Cole S., Alizon M.;  
RT "Nucleotide sequence of the AIDS virus, LAV.#";  
RL Cell 40:9-17(1985).  
RN [2]  
RP SEQUENCE FROM N.A. (CLONE PNL4-3).  
RA Buckler C.E., Buckler-White A.J., Willey R.L., McCoy J.;  
RL Submitted (JUN-1988) to the EMBL/GenBank/DBJ databases.  
CC -!- FUNCTION: TRANSCRIPTIONAL REGULATOR THAT ACTS BY BINDING TO THE  
CC TRANS-ACTIVATING RESPONSIVE SEQUENCE (TAR) RNA ELEMENT AND  
CC ACTIVATES TRANSCRIPTION INITIATION AND/OR ELONGATION FROM THE LTR  
CC PROMOTER.  
CC -!- SUBUNIT: BINDS CYCLIN T1 (BY SIMILARITY).  
CC -!- SUBCELLULAR LOCATION: NUCLEAR; NUCLEOLAR.  
CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
CC the European Bioinformatics Institute. There are no restrictions on its  
CC use by non-profit institutions as long as its content is in no way  
CC modified and this statement is not removed. Usage by and for commercial  
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
CC  
CC EMBL: K02013; AAB59745.1; -;  
CC EMBL: M19921; AAB44985.1; -;  
CC HIV: K02013; TAT8BRU.  
CC HIV: M19921; TAT8NL43.  
CC InterPro: IPR001831; HIV\_Tat.  
CC Pfam: PF00539; Tat. 1.  
CC PRINTS: PR00055; HIVTATDOMAIN.  
CC Transcription regulation; Activator; RNA-binding; Nuclear protein;  
KW AIDS.  
FT VARIANT 24 24 T -> N (IN CLONE PNL4-3).  
FT VARIANT 39 39 T -> M (IN CLONE PNL4-3).  
FT VARIANT 58 61 PPQG -> AHON (IN CLONE PNL4-3).  
FT VARIANT 67 67 V -> A (IN CLONE PNL4-3).  
FT VARIANT 77 77 P -> S (IN CLONE PNL4-3).  
SEQUENCE 86 AA; 9769 MW; 9B1BA915FAF8A14 CRC64;

Query Match 20.8%; Score 470; DB 1; Length 86;  
Best Local Similarity 96.5%; Pred. No. 1.9e-32;  
Matches 82; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

QY 318 EPVDPRLPEWKHPGSPKTCACNYCKCCFHCQVCFITKALGISYGRKKRRRRPPQG 377  
|||||  
Db 2 EPVDPRLPEWKHPGSPKTCACNYCKCCFHCQVCFITKALGISYGRKKRRRRPPQG 61  
|||||

QY 378 SOTHQVSLSKQPTSQSRGDPGPK 402  
|||||  
Db 62 SOTHQVSLSKQPTSQSRGDPGPK 86  
|||||

## RESULT 30

TAT\_HV1BR STANDARD: PRT: 102 AA.  
ID TAT\_HV1BR  
AC P05908;  
DT 01-NOV-1988 (Rel. 09, Created)  
DT 01-NOV-1988 (Rel. 09, Last sequence update)  
DT 16-OCT-2001 (Rel. 40, Last annotation update)  
DE TAT protein (Transactivating regulatory protein).  
GN TAT.  
OS Human immunodeficiency virus type 1 (RF/HAT isolate) (HIV-1).  
OC Viruses; Retroviral viruses; Retroviridae; Lentivirus.  
OX NCBI\_TaxID=11701;  
RN [1]  
RP SEQUENCE FROM N.A.  
RX MEDLINE=86218077; PubMed=2423250;  
RA Starcich B.R., Hahn B.H., Shaw G.M., McNeely P.D., Modrow S.,  
RA Wolf H., Parks E.S., Parks W.P., Josephs S.F., Gallo R.C.,  
RA Wong-Staal F.;  
RT "Identification and characterization of conserved and variable  
RT regions in the envelope gene of HTLV-III/LAV, the retrovirus of  
RT AIDS.";  
RL Cell 45:637-648(1986).  
CC -!- FUNCTION: TRANSCRIPTIONAL REGULATOR THAT ACTS BY BINDING TO THE  
CC TRANS-ACTIVATING RESPONSIVE SEQUENCE (TAR) RNA ELEMENT AND  
CC ACTIVATES TRANSCRIPTION INITIATION AND/OR ELONGATION FROM THE LTR  
CC PROMOTER.  
CC -!- SUBUNIT: BINDS CYCLIN T1 (BY SIMILARITY).  
CC -!- SUBCELLULAR LOCATION: NUCLEAR; NUCLEOLAR.  
CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
CC the European Bioinformatics Institute. There are no restrictions on its  
CC use by non-profit institutions as long as its content is in no way  
CC modified and this statement is not removed. Usage by and for commercial  
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
CC  
CC EMBL: M17451; AAB45050.1; -;  
CC HIV: M17451; TAT8NF.  
CC InterPro: IPR001831; HIV\_Tat.  
CC Pfam: PF00539; Tat. 1.  
CC PRINTS: PR00055; HIVTATDOMAIN.  
CC Transcription regulation; Activator; RNA-binding; Nuclear protein;  
KW AIDS.  
SQ SEQUENCE 102 AA; 11538 MW; 3EACFB843C5195BD CRC64;

Query Match 20.4%; Score 461; DB 1; Length 102;  
Best Local Similarity 91.9%; Pred. No. 1.3e-31;  
Matches 79; Conservative 3; Mismatches 4; Indels 0; Gaps 0;

QY 318 EPVDPRLPEWKHPGSPKTCACNYCKCCFHCQVCFITKALGISYGRKKRRRRPPQG 377  
|||||  
Db 2 EPVDPRLPEWKHPGSPKTCACNYCKCCFHCQVCFITKALGISYGRKKRRRRPPQG 61  
|||||  
QY 378 SOTHQVSLSKQPTSQSRGDPGPK 403  
|||||  
Db 62 SOTHQVSLSKQPTSQSRGDPGPK 87  
|||||

Search completed: August 26, 2002, 08:15:42  
Job time: 346 sec

GenCore version 4.5  
Copyright (c) 1993 - 2000 Compugen Ltd.

OM protein - protein search, using sw model

Run on: August 26, 2002, 08:14:55 ; Search time 86.53 Seconds  
(without alignments)  
821.691 Million cell updates/sec

Title: us-09-509-239-17  
Perfect score: 2255  
Sequence: 1 CSHSSNMANTOMKSDKIII.....QSRGDTGPKTSGHHHHH 411

Scoring table:  
BLOSUM62  
Gapop 10.0 , Gapext 0.5

Searched: 562222 seqs, 172994929 residues

Total number of hits satisfying chosen parameters: 562222

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : SPTREMBL\_19.\*

1: sp\_archaea.\*  
2: sp\_bacteria.\*  
3: sp\_fungi.\*  
4: sp\_human.\*  
5: sp\_invertebrate.\*  
6: sp\_mammal.\*  
7: sp\_mhc.\*  
8: sp\_organelle.\*  
9: sp\_phase.\*  
10: sp\_plant.\*  
11: sp\_rodent.\*  
12: sp\_virus.\*  
13: sp\_vertebrate.\*  
14: sp\_unclassified.\*  
15: sp\_rvirus.\*  
16: sp\_bacteriap.\*  
17: sp\_archaeap.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

#### SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	1129	50.1	206	15 Q9WM24	Q9wm24 human immun
2	1124	49.8	206	15 Q9WLM4	Q9wlm4 human immun
3	1119	49.6	206	15 Q40177	Q40177 human immun
4	1115	49.4	206	15 Q90V07	Q90v07 human immun
5	1108	49.1	206	15 Q9PXW9	Q9pxw9 human immun
6	1105	49.0	206	15 Q9PXW9	Q9pxw9 human immun
7	1104	49.0	206	15 Q9WM30	Q9wm30 human immun
8	1101	48.8	206	15 Q9PXW8	Q9pxw8 human immun
9	1096	48.6	206	15 Q90179	Q90179 human immun
10	1096	48.6	206	15 Q85368	Q85368 human immun
11	1096	48.6	206	15 Q78244	Q78244 human immun
12	1092	48.4	206	15 Q9WLM7	Q9wlm7 human immun
13	1090	48.3	206	15 Q74905	Q74905 human immun
14	1087	48.2	206	15 Q89561	Q89561 human immun
15	1085	48.1	206	15 Q74913	Q74913 human immun
16	1080	47.9	206	15 Q74917	Q74917 human immun

#### ALIGNMENTS

RESULT 1

Q9WM24 ID Q9WM24 PRELIMINARY; PRT; 206 AA.  
AC Q9WM24;  
DT 01-NOV-1999 (Tremblrel. 12, Created)  
DT 01-NOV-1999 (Tremblrel. 12, Last sequence update)  
DT 01-DEC-2001 (Tremblrel. 19, Last annotation update)  
DE NEGATIVE FACTOR (F-PROTEIN) (27 KDA PROTEIN).  
GN NEF.  
OS Human immunodeficiency virus type 1.  
OC Viruses; Retroid viruses; Retroviridae; Lentivirus.  
OX NCBI\_taxID=11676;  
RN [1]  
RP SEQUENCE FROM N.A.  
RC STRAIN=12-IT;  
RA Visco Comandini U., Yun Z., Paganelli R., Orlandi P., Salotti A.,  
RA Johansson B., Vahine A., Sonnerborg A.;  
RT "HIV-1 nef mutations and clinical long-term non progression: a  
RT molecular epidemiology study";  
RL Submitted (JUN-1997) to the EMBL/GenBank/DBJ databases.  
CC -1- SIMILARITY: NEF HAS GTPASE, GTP-BINDING AND AUTOPHOSPHORYLATING  
CC ACTIVITIES, IT SEEMS TO DOWN-REGULATE THE CD4(T4) ANTIGEN.  
DR EMBL; AF011480; AAD01458.1; -  
DR HSSP; P03406; 1EPN.  
DR InterPro; IPR001558; F-protein.  
DR Pfam; PF00469; F-protein; 1.  
KW AIDS; GTP-binding; Myristate.  
SQ SEQUENCE 206 AA; 23328 MW; FCC69458158FIA03 CRC64;

Query Match 50.1%; Score 1129; DB 15; Length 206;  
Best Local Similarity 99.5%; Pred. No. 2.7e-83;  
Matches 205; Conservative 1; Mismatches 0; Indels 0; Gaps 0;  
QY 110 MCGKWSKSVVGVVREMRRAEPAADGVGAASRDLEKHGAITSNTAATNAACAWLEA 169  
DB 1 MCGKWSKSVVGVVREMRRAEPAADGVGAASRDLEKHGAITSNTAATNAACAWLEA 60  
QY 170 QEEVEVGFPVTPQVPLRPMTYKAAVDLSHFLKEKGLGLIHSQRQDILDLIWYHTQGY 229



Query Match 49.4%; Score 1115; DB 15; Length 206;  
 Best Local Similarity 98.1%; Pred. No. 3.7e-82;  
 Matches 202; Conservative 1; Mismatches 3; Indels 0; Gaps 0;

QY 110 MGKSKSVVGVGPTVRRMRRAEPAADGGAASRDLEKHGAITSNTAATNAACAWLEA 169  
 DB 1 MGKSKSVVGVGPTVRRMRRAEPAADGGAASRDLEKHGAITSNTAATNAACAWLEA 60

QY 170 QEEVEGFPVTPQVPLRPMYTYKAADLSHFLKEKGGLEGLIHSQRQDILDLMWYHTQY 229  
 DB 61 QEEVEGFPVTPQVPLRPMYTYKAADLSHFLKEKGGLEGLIHSQRQDILDLMWYHTQY 120

QY 230 FPDWQNTGPGVRYPLTFCWCYKLVPEPDKVEEANKGENTSLHPVSLHGMDPEREV 289  
 DB 121 FPDWQNTGPGVRYPLTFCWCYKLVPEPDKVEEANKGENTSLHPVSLHGMDPEREV 180

QY 290 LEWRDLSRLAFHHVARELHPEYFKNC 315  
 DB 181 LEWRDLSRLAFHHVARELHPEYFKNC 206

RESULT 5  
 Q9PW9 PRELIMINARY; PRT; 206 AA.  
 AC Q9PW9;  
 DT 01-MAY-2000 (Tremblrel. 13, Created)  
 DT 01-MAY-2000 (Tremblrel. 13, Last sequence update)  
 DT 01-DEC-2001 (Tremblrel. 19, Last annotation update)  
 DE NEGATIVE FACTOR (F-PROTEIN) (27 KDA PROTEIN).  
 GN NEF.  
 OS Human immunodeficiency virus type 1.  
 OC Viruses; Retroid viruses; Retroviridae; Lentivirus.  
 OX NCBI\_TaxID=11676;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RA Kohleisen B., Neumann M., Herrmann R., Brack-Werner R., Krohn K.J.,  
 RA Ovod V., Ranki A., Erle V.;  
 RT "Cellular localization of Nef expressed in persistently HIV-1-infected  
 RT low-producer astrocytes";  
 RL AIDS 6:1427-1436(1992).  
 CC -1- SIMILARITY: NEF HAS GTPASE, GTP-BINDING AND AUTOPHOSPHORYLATING  
 CC ACTIVITIES, IT SEEMS TO DOWN-REGULATE THE CD4(T4) ANTIGEN.  
 DR HSP; P03406; IEFN.  
 DR InterPro; IPR001558; F-protein.  
 DR Pfam; PF00469; F-protein; 1.  
 KW AIDS; GTP-binding; Myristate.  
 SQ SEQUENCE 206 AA; 23454 MW; 65D5DDE4FB748072 CRC64;

Query Match 49.1%; Score 1108; DB 15; Length 206;  
 Best Local Similarity 97.1%; Pred. No. 1.3e-81;  
 Matches 200; Conservative 3; Mismatches 3; Indels 0; Gaps 0;

QY 110 MGKSKSVVGVGPTVRRMRRAEPAADGGAASRDLEKHGAITSNTAATNAACAWLEA 169  
 DB 1 MGKSKSVVGVGPTVRRMRRAEPAADGGAASRDLEKHGAITSNTAATNAACAWLEA 60

QY 170 QEEVEGFPVTPQVPLRPMYTYKAADLSHFLKEKGGLEGLIHSQRQDILDLMWYHTQY 229  
 DB 61 QEEVEGFPVTPQVPLRPMYTYKAADLSHFLKEKGGLEGLIHSQRQDILDLMWYHTQY 120

QY 230 FPDWQNTGPGVRYPLTFCWCYKLVPEPDKVEEANKGENTSLHPVSLHGMDPEREV 289  
 DB 121 FPDWQNTGPGVRYPLTFCWCYKLVPEPDKVEEANKGENTSLHPVSLHGMDPEREV 180

QY 290 LEWRDLSRLAFHHVARELHPEYFKNC 315  
 DB 181 LEWRDLSRLAFHHVARELHPEYFKNC 206

Query Match 49.1%; Score 1108; DB 15; Length 206;  
 Best Local Similarity 97.1%; Pred. No. 1.3e-81;  
 Matches 200; Conservative 3; Mismatches 3; Indels 0; Gaps 0;

QY 110 MGKSKSVVGVGPTVRRMRRAEPAADGGAASRDLEKHGAITSNTAATNAACAWLEA 169  
 DB 1 MGKSKSVVGVGPTVRRMRRAEPAADGGAASRDLEKHGAITSNTAATNAACAWLEA 60

QY 170 QEEVEGFPVTPQVPLRPMYTYKAADLSHFLKEKGGLEGLIHSQRQDILDLMWYHTQY 229  
 DB 61 QEEVEGFPVTPQVPLRPMYTYKAADLSHFLKEKGGLEGLIHSQRQDILDLMWYHTQY 120

QY 230 FPDWQNTGPGVRYPLTFCWCYKLVPEPDKVEEANKGENTSLHPVSLHGMDPEREV 289  
 DB 121 FPDWQNTGPGVRYPLTFCWCYKLVPEPDKVEEANKGENTSLHPVSLHGMDPEREV 180

QY 290 LEWRDLSRLAFHHVARELHPEYFKNC 315  
 DB 181 LEWRDLSRLAFHHVARELHPEYFKNC 206

## RESULT 6

Q9WM16 PRELIMINARY; PRT; 206 AA.  
 AC Q9WM16;  
 DT 01-NOV-1999 (Tremblrel. 12, Created)  
 DT 01-NOV-1999 (Tremblrel. 12, Last sequence update)  
 DT 01-DEC-2001 (Tremblrel. 19, Last annotation update)  
 DE NEGATIVE FACTOR (F-PROTEIN) (27 KDA PROTEIN).  
 GN NEF.  
 OS Human immunodeficiency virus type 1.  
 OC Viruses; Retroid viruses; Retroviridae; Lentivirus.  
 OX NCBI\_TaxID=11676;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN-7-IT;  
 RA Visco Comandini U., Yun Z., Paganelli R., Orlandi P., Salotti A.,  
 RA Johansson B., Vahne A., Sonnerborg A.;  
 RT "HIV-1 nef mutations and clinical long-term non progression: a  
 RT molecular epidemiology study";  
 RL Submitted (JUN-1997) to the EMBL/GenBank/DBJ databases.  
 CC -1- SIMILARITY: NEF HAS GTPASE, GTP-BINDING AND AUTOPHOSPHORYLATING  
 CC ACTIVITIES, IT SEEMS TO DOWN-REGULATE THE CD4(T4) ANTIGEN.  
 DR EMBL; AF011494; AAD01472.1;  
 DR HSP; P03406; IEFN.  
 DR InterPro; IPR001558; F-protein.  
 DR Pfam; PF00469; F-protein; 1.  
 KW AIDS; GTP-binding; Myristate.  
 SQ SEQUENCE 206 AA; 23333 MW; 2EBF2A6A3ECAF5EA CRC64;

Query Match 49.0%; Score 1105; DB 15; Length 206;  
 Best Local Similarity 97.1%; Pred. No. 2.3e-81;  
 Matches 200; Conservative 2; Mismatches 4; Indels 0; Gaps 0;

QY 110 MGKSKSVVGVGPTVRRMRRAEPAADGGAASRDLEKHGAITSNTAATNAACAWLEA 169  
 DB 1 MGKSKSVVGVGPTVRRMRRAEPAADGGAASRDLEKHGAITSNTAATNAACAWLEA 60

QY 170 QEEVEGFPVTPQVPLRPMYTYKAADLSHFLKEKGGLEGLIHSQRQDILDLMWYHTQY 229  
 DB 61 QEEVEGFPVTPQVPLRPMYTYKAADLSHFLKEKGGLEGLIHSQRQDILDLMWYHTQY 120

QY 230 FPDWQNTGPGVRYPLTFCWCYKLVPEPDKVEEANKGENTSLHPVSLHGMDPEREV 289  
 DB 121 FPDWQNTGPGVRYPLTFCWCYKLVPEPDKVEEANKGENTSLHPVSLHGMDPEREV 180

QY 290 LEWRDLSRLAFHHVARELHPEYFKNC 315  
 DB 181 LEWRDLSRLAFHHVARELHPEYFKNC 206

## RESULT 7

Q9WM30 PRELIMINARY; PRT; 206 AA.  
 AC Q9WM30;  
 DT 01-NOV-1999 (Tremblrel. 12, Created)  
 DT 01-NOV-1999 (Tremblrel. 12, Last sequence update)  
 DT 01-DEC-2001 (Tremblrel. 19, Last annotation update)  
 DE NEGATIVE FACTOR (F-PROTEIN) (27 KDA PROTEIN).  
 GN NEF.  
 OS Human immunodeficiency virus type 1.  
 OC Viruses; Retroid viruses; Retroviridae; Lentivirus.  
 OX NCBI\_TaxID=11676;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN-8-IT;  
 RA Visco Comandini U., Yun Z., Paganelli R., Orlandi P., Salotti A.,  
 RA Johansson B., Vahne A., Sonnerborg A.;  
 RT "HIV-1 nef mutations and clinical long-term non progression: a  
 RT molecular epidemiology study";  
 RL Submitted (JUN-1997) to the EMBL/GenBank/DBJ databases.  
 CC -1- SIMILARITY: NEF HAS GTPASE, GTP-BINDING AND AUTOPHOSPHORYLATING  
 CC ACTIVITIES, IT SEEMS TO DOWN-REGULATE THE CD4(T4) ANTIGEN.

```
DR EMBL; AF011469; AAD01447.1; -.
DR HSSP; P03406; 1EFN.
DR InterPro; IPR001558; F-protein.
DR Pfam; PF00469; F-protein; 1.
KW AIDS; GTP-binding; Myristate.
SQ SEQUENCE 206 AA; 23345 MW; 0ED69927C2E03BB6 CRC64;

Query Match 49.0%; Score 1104; DB 15; Length 206;
Best Local Similarity 97.1%; Pred. No. 2.8e-81;
Matches 200; Conservative 1; Mismatches 5; Indels 0; Gaps 0;

QY 110 MGGKSKSSVGVGPTVVRMRRAEPAADGVGAASRDLEKHGAITSSNTAATNAACAWLEA 169
DB 1 MGGKSKSSVIGWPAVRMRRAEPAADGVGAASRDLEKHGAITSSNTAANNAACAWLEA 60

QY 170 QEEVEGFPVTPQVPLRPMTYKAADVLSHFLKEKGLGEGLIHSQRQDILDWIYHTQGY 229
DB 61 QEEVEGFPVTPQVPLRPMTYKAADVLSHFLKEKGLGGLXGLIHSQRQDILDWIYHTQGY 120

QY 230 FPDQNYTPGQVRYPLTFGWCYKLVPEPDKVEBANKGENTSLHPVSLHGMDPPERV 289
DB 121 FPDQNYTPGQVRYPLTFGWCYKLVPEPDKVEBANKGENTSLHPVSLHGMDPPERV 180

QY 290 LEWRDLSRLAFHHVARELHPEYFKNC 315
DB 181 LEWRDLSRLAFHHVARELHPEYFKNC 206

RESULT 8
Q9PXW8 PRELIMINARY; PRT; 206 AA.
AC Q9PXW8;
DT 01-MAY-2000 (TrEMBLrel. 13, Created)
DT 01-MAY-2000 (TrEMBLrel. 13, Last sequence update)
DT 01-DEC-2001 (TrEMBLrel. 19, Last annotation update)
DE NEGATIVE FACTOR (F-PROTEIN) (27 KDA PROTEIN).
GN NEF.
OS Human immunodeficiency virus type 1.
OC Viruses; Retroid viruses; Retroviridae; Lentivirus.
OX NCBI_TaxID=11676;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=93152025; PubMed=1301062;
RT Kohleisen B., Neumann M., Herrmann R., Brack-Werner R., Krohn K.J.,
RT Ovod V., Ranki A., Erfle V.;
RT "Cellular localization of Nef expressed in persistently HIV-1-infected
RT low-producer astrocytes.";
RL AIDS 6:1427-1436(1992).
CC -1- SIMILARITY: NEF HAS GTPASE, GTP-BINDING AND AUTOPHOSPHORYLATING
CC ACTIVITIES. IT SEEMS TO DOWN-REGULATE THE CD4(74) ANTIGEN.
DR HSSP; P03406; 1EFN.
DR InterPro; IPR001558; F-protein.
DR Pfam; PF00469; F-protein; 1.
KW AIDS; GTP-binding; Myristate.
SQ SEQUENCE 206 AA; 23439 MW; 8478CF194D39D5B9 CRC64;

Query Match 48.8%; Score 1101; DB 15; Length 206;
Best Local Similarity 96.1%; Pred. No. 4.9e-81;
Matches 198; Conservative 4; Mismatches 4; Indels 0; Gaps 0;

QY 110 MGGKSKSSVGVGPTVVRMRRAEPAADGVGAASRDLEKHGAITSSNTAATNAACAWLEA 169
DB 1 MGGKSKSSVIGWPAVRMRRAEPAADGVGAASRDLEKHGAITSSNTAANNAACAWLEA 60

QY 170 QEEVEGFPVTPQVPLRPMTYKAADVLSHFLKEKGLGEGLIHSQRQDILDWIYHTQGY 229
DB 61 QEEVEGFPVTPQVPLRPMTYKAADVLSHFLKEKGLGGLIHSQRQDILDWIYHTQGY 120

QY 230 FPDQNYTPGQVRYPLTFGWCYKLVPEPDKVEBANKGENTSLHPVSLHGMDPPERV 289
DB 121 FPDQNYTPGQVRYPLTFGWCYKLVPEPDKVEBANKGENTSLHPVSLHGMDPPERV 180
```

```
QY 290 LEWRDLSRLAFHHVARELHPEYFKNC 315
DB 181 LEWRDLSRLAFHHVARELHPEYFKNC 206

RESULT 9
Q90179 PRELIMINARY; PRT; 206 AA.
AC Q90179;
DT 01-NOV-1998 (TrEMBLrel. 08, Created)
DT 01-NOV-1998 (TrEMBLrel. 08, Last sequence update)
DT 01-DEC-2001 (TrEMBLrel. 19, Last annotation update)
DE NEGATIVE FACTOR (F-PROTEIN) (27 KDA PROTEIN).
GN NEF.
OS Human immunodeficiency virus type 1.
OC Viruses; Retroid viruses; Retroviridae; Lentivirus.
OX NCBI_TaxID=11676;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=95074930; PubMed=7983770;
RT Fang H., Pincus S.H.;
RT "Unique insertion sequence and pattern of CD4 expression in variants
RT selected with immunotoxins from human immunodeficiency virus type 1-
RT infected T cells.";
RL J. Virol. 69:75-81(1995).
RN [2]
RP SEQUENCE FROM N.A.
RA Fang H., Pincus S.H.;
RT "Spontaneous activation of human immunodeficiency virus type 1 in an
RT immunotoxin-resistant variant T cell line.";
RT Submitted (JUN-1998) to the EMBL/GenBank/DDSI databases.
CC -1- SIMILARITY: NEF HAS GTPASE, GTP-BINDING AND AUTOPHOSPHORYLATING
CC ACTIVITIES. IT SEEMS TO DOWN-REGULATE THE CD4(74) ANTIGEN.
DR EMBL; AF070521; AAC28453.1; -.
DR HSSP; P03406; 1EFN.
DR InterPro; IPR001558; F-protein.
DR Pfam; PF00469; F-protein; 1.
KW AIDS; GTP-binding; Myristate.
SQ SEQUENCE 206 AA; 23334 MW; 3E2B9C4017FDC6A8 CRC64;

Query Match 48.6%; Score 1096; DB 15; Length 206;
Best Local Similarity 95.6%; Pred. No. 1.2e-80;
Matches 197; Conservative 4; Mismatches 5; Indels 0; Gaps 0;

QY 110 MGGKSKSSVGVGPTVVRMRRAEPAADGVGAASRDLEKHGAITSSNTAATNAACAWLEA 169
DB 1 MGGKSKSSVIGWPAVRMRRAEPAADGVGAASRDLEKHGAITSSNTAANNAACAWLEA 60

QY 170 QEEVEGFPVTPQVPLRPMTYKAADVLSHFLKEKGLGEGLIHSQRQDILDWIYHTQGY 229
DB 61 QEEVEGFPVTPQVPLRPMTYKAADVLSHFLKEKGLGGLIHSQRQDILDWIYHTQGY 120

QY 230 FPDQNYTPGQVRYPLTFGWCYKLVPEPDKVEEANKGENTSLHPVSLHGMDPPERV 289
DB 121 FPDQNYTPGQVRYPLTFGWCYKLVPEPDKVEEANKGENTSLHPVSLHGMDPPERV 180

QY 290 LEWRDLSRLAFHHVARELHPEYFKNC 315
DB 181 LEWRDLSRLAFHHVARELHPEYFKNC 206

RESULT 10
Q85588 PRELIMINARY; PRT; 206 AA.
AC Q85588;
DT 01-NOV-1996 (TrEMBLrel. 01, Created)
DT 01-NOV-1996 (TrEMBLrel. 01, Last sequence update)
DT 01-DEC-2001 (TrEMBLrel. 19, Last annotation update)
DE NEGATIVE FACTOR (F-PROTEIN) (27 KDA PROTEIN).
GN NEF.
OS Aids-associated retrovirus.
```



```

OC Viruses; Retroviral viruses; Retroviridae.
OX NCBI_TaxID=11966;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=86067226; PubMed=2999715;
RA Ratner L., Starcich B., Josephs S.F., Hahn B.H., Reddy E.P., Arya S.K.,
RA Liyak K.J., Pettaway S.R., Pearson M.L., Haseltine W.A., Arva S.K.,
RA Wong-staal F.;
RT "Polymorphism of the 3' open reading frame of the virus associated
RT with the acquired immune deficiency syndrome, human T-lymphotropic
RT virus type III.";
RL Nucleic Acids Res. 13:8219-8229(1985).
CC -1- SIMILARITY: NEF HAS GTPASE, GTP-BINDING AND AUTOPHOSPHORYLATING
CC ACTIVITIES, IT SEEMS TO DOWN-REGULATE THE CD4(T4) ANTIGEN.
DR EMBL; X03188; CAA26947.1; -.
DR HSSP; P03406; 1EFN.
DR InterPro; IPR001558; F-protein.
DR Pfam; PF00469; F-protein; 1.
KW AIDS; GTP-binding; Myristate.
SQ SEQUENCE 206 AA; 23419 MW; 3B25EB332A479A46 CRC64;

Query Match 48.6%; Score 1096; DB 15; Length 206;
Best Local Similarity 96.6%; Pred. No. 1.2e-80;
Matches 199; Conservative 4; Mismatches 3; Indels 0; Gaps 0;

Qy 110 MGGKWSKSSVVGWPTVRMRRAEPAADGVGAASRDLEKHGAITSNTAATNAACAWLEA 169
Db 1 MGGKWSKSSVVGWPAVRMRRAEPAADGVGAASRDLEKHGAITSNTAANNAACAWLEA 60

Qy 170 QEEEVGFPVTPQVPLRPMTYKAAVDLSHFLKEKGLGLIHSQRQDILDWIYHTQGY 229
Db 61 QEEEVGFPVTPQVPLRPMTYKAAVDLSHFLKEKGLGLIHSQRQDILDWIYHTQGY 120

Qy 230 FPDQNYTPGCGVRYPLTFGWCYKLVPEPKVEANKGENTSLHPVSLHGMDPPEV 289
Db 121 FPDQNYTPGCGIRYPLTFGWCYKLVPEPKVEANKGENTSLHPVSLHGMDPPEV 180

Qy 290 LEWRDLSRLAFHHVARELHPEYFKNC 315
Db 181 LEWRDLSRLAFHHVARELHPEYFKNC 206

RESULT 11
Q78244 PRELIMINARY; PRT; 206 AA.
ID Q78244;
AC Q78244;
DT 01-NOV-1996 (TrEMBLrel. 01, Created)
DT 01-NOV-1996 (TrEMBLrel. 01, Last sequence update)
DT 01-DEC-2001 (TrEMBLrel. 19, Last annotation update)
DE NEGATIVE FACTOR (F-PROTEIN) (27 KDA PROTEIN).
GN NEF.
OS Human immunodeficiency virus type 1.
OC Viruses; Retroviral viruses; Retroviridae; Lentivirus.
OX NCBI_TaxID=11676;
RN [1]
RP SEQUENCE FROM N.A.
RA Carlini F., Federic M., Equestre M., Ricci S., Ratti G., Zibai Q.,
RA Verani P., Rossi G.B.;
RT "Sequence analysis of HIV-1 proviral DNA from a non producer
RT chronically infected HUT-78 cellular clone.";
RL J. Viral Diseases 1:40-55(1992).
RN [2]
RP SEQUENCE FROM N.A.
RX MEDLINE=89352106; PubMed=2765297;
RA Federico M., Titti F., Butto S., Orecchia A., Carlini F., Taddeo B.,
RA Macchi B., Mangiano N., Verani P., Rossi G.;
RT "Biological and molecular characterization of producer and non
RT producer clones from HUT-78 infected with a patient HIV isolate.";
RL AIDS Res. Hum. Retroviruses 5:385-396(1989).
RN [3]
RP SEQUENCE FROM N.A.
RA Titti F., Federico M., Butto S., Orecchia A., Carlini F., Taddeo B.,

```

```

RA Borsetti A., Saggio I., Verani P., Rossi G.;
RT "Variability of HIV-1 virus: characteristics of an infected but not
RT productive clone.";
RL Int. J. Immunopharmacol. 3:17-23(1990).
CC -1- SIMILARITY: NEF HAS GTPASE, GTP-BINDING AND AUTOPHOSPHORYLATING
CC ACTIVITIES, IT SEEMS TO DOWN-REGULATE THE CD4(T4) ANTIGEN.
DR EMBL; Z11530; CAA77629.1; -.
DR HSSP; P03406; 1EFN.
DR InterPro; IPR001558; F-protein.
DR Pfam; PF00469; F-protein; 1.
KW AIDS; GTP-binding; Myristate.
SQ SEQUENCE 206 AA; 23398 MW; A31CC7244BBB2B43 CRC64;

Query Match 48.6%; Score 1096; DB 15; Length 206;
Best Local Similarity 96.6%; Pred. No. 1.2e-80;
Matches 199; Conservative 3; Mismatches 4; Indels 0; Gaps 0;

Qy 110 MGGKWSKSSVVGWPTVRMRRAEPAADGVGAASRDLEKHGAITSNTAATNAACAWLEA 169
Db 1 MGGKWSKSSVVGWPAVRMRRAEPAADGVGAASRDLEKHGAITSNTAATNADCAWLEA 60

Qy 170 QEEEVGFPVTPQVPLRPMTYKAAVDLSHFLKEKGLGLIHSQRQDILDWIYHTQGY 229
Db 61 QEEEVGFPVTPQVPLRPMTYKAAVDLSHFLKEKGLGLIHSQRQDILDWIYHTQGY 120

Qy 230 FPDQNYTPGCGVRYPLTFGWCYKLVPEPKVEANKGENTSLHPVSLHGMDPPEV 289
Db 121 FPDQNYTPGCGIRYPLTFGWCYKLVPEPKVEANKGENTSLHPVSLHGMDPPEV 180

Qy 290 LEWRDLSRLAFHHVARELHPEYFKNC 315
Db 181 LEWRDLSRLAFHHVARELHPEYFKNC 206

RESULT 12
Q9WLM7 PRELIMINARY; PRT; 206 AA.
ID Q9WLM7;
AC Q9WLM7;
DT 01-NOV-1999 (TrEMBLrel. 12, Created)
DT 01-NOV-1999 (TrEMBLrel. 12, Last sequence update)
DT 01-DEC-2001 (TrEMBLrel. 19, Last annotation update)
DE NEGATIVE FACTOR (F-PROTEIN) (27 KDA PROTEIN).
GN NEF.
OS Human immunodeficiency virus type 1.
OC Viruses; Retroviral viruses; Retroviridae; Lentivirus.
OX NCBI_TaxID=11676;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=18-SW;
RA Visco Comandini U., Yun Z., Paganelli R., Orlandi P., Salotti A.,
RA Johanson B., Vahne A., Sonnerborg A.;
RT "HIV-1 nef mutations and clinical long-term non progression: a
RT molecular epidemiology study.";
RL Submitted (FEB-1998) to the EMBL/GenBank/DBJ databases.
CC -1- SIMILARITY: NEF HAS GTPASE, GTP-BINDING AND AUTOPHOSPHORYLATING
CC ACTIVITIES, IT SEEMS TO DOWN-REGULATE THE CD4(T4) ANTIGEN.
DR EMBL; AF047084; AAD02458.1; -.
DR HSSP; P03406; 1EFN.
DR InterPro; IPR001558; F-protein.
DR Pfam; PF00469; F-protein; 1.
KW AIDS; GTP-binding; Myristate.
SQ SEQUENCE 206 AA; 23512 MW; E09E3BEF828A83C0 CRC64;

```

```

Query Match 48.4%; Score 1092; DB 15; Length 206;
Best Local Similarity 95.6%; Pred. No. 2.6e-80;
Matches 197; Conservative 7; Mismatches 2; Indels 0; Gaps 0;

Qy 110 MGGKWSKSSVVGWPTVRMRRAEPAADGVGAASRDLEKHGAITSNTAATNAACAWLEA 169
Db 1 MGGKWSKSSVVGWPTVRMRRAEPAADGVGAASRDLEKHGAITSNTAATNACAWLEA 60

```

QY 170 QEEVEGFPVTPQVPLRPMTYKAAVDLSHFLKEKGLGSLHSQRQDILDLDWIYHTQY 229  
|||||  
DB 61 QEEVEGFPVTPQVPLRPMTYKAAVDLSHFLKEKGLGSLHSQRQDILDLDWIYHTQY 120  
|||||  
QY 230 FPDQNYTPGCVRYPLTFGWCYKLVPEPKVEANKGENTSLHPVSLHGMDDPEREV 289  
|||||  
DB 121 FPDQNYTPGCVRYPLTFGWCYKLVPEPKVEANKGENTSLHPVSLHGMDDPEREV 180  
|||||  
QY 290 LEWRDLSRLAFHHVARELHPEYFKNC 315  
|||||  
DB 181 LEWRDLSRLAFHHVARELHPEYFKNC 206  
|||||  
RESULT 13  
Q74905 PRELIMINARY; PRT; 206 AA.  
ID Q74905;  
AC Q74905;  
DT 01-NOV-1996 (TremBLrel. 01, Created)  
DT 01-NOV-1996 (TremBLrel. 01, Last sequence update)  
DT 01-DEC-2001 (TremBLrel. 19, Last annotation update)  
DE NEGATIVE FACTOR (F-PROTEIN) (27 KDA PROTEIN).  
GN NEF.  
OS Human immunodeficiency virus type 1.  
OC Viruses; Retroviral viruses; Retroviridae; Lentivirus.  
OX NCBI\_TaxID=11676;  
RN [1]  
RP SEQUENCE FROM N.A.  
RX MEDLINE=96400183; PubMed=8806559;  
RA Ratner L., Joseph T., Bandres J., Ghosh S., Heyden N.V., Templeton A.,  
Hahn B., Powderly W., Arens M.,  
RT "Sequence heterogeneity of Nef transcripts in HIV-1-infected subjects  
at different stages of disease."  
RL Virology 223:245-250(1996).  
CC -I- SIMILARITY: NEF HAS GTPASE, GTP-BINDING AND AUTOPHOSPHORYLATING  
ACTIVITIES, IT SEEMS TO DOWN-REGULATE THE CD4(T4) ANTIGEN.  
DR EMBL; U4443; AAB38195.1; -;  
DR HSP; Q70627; 2NEF.  
DR InterPro: IPR001558; F-protein.  
DR Pfam: PF00469; F-protein; 1.  
DR AIDS; GTP-binding; Myristate.  
KW AIDS; GTP-binding; Myristate.  
SQ SEQUENCE 206 AA; 23617 MW; 5CAAB09C4730C38A CRC64;

Query Match 48.3%; Score 1090; DB 15; Length 206;  
Best Local Similarity 95.1%; Pred. No. 3.8e-80;  
Matches 196; Conservative 6; Mismatches 4; Indels 0; Gaps 0;  
QY 110 MGGKWSKSSVVGWPTVRMRRAEPAADGVGAASRDLEKKGSLHSQRQDILDLDWIYHTQY 169  
|||||  
DB 1 MGGKWSKSSVVGWPTVRMRRAEPAADGVGAASRDLEKKGSLHSQRQDILDLDWIYHTQY 60  
|||||  
QY 170 QEEVEGFPVTPQVPLRPMTYKAAVDLSHFLKEKGLGSLHSQRQDILDLDWIYHTQY 229  
|||||  
DB 61 QEEVEGFPVTPQVPLRPMTYKAAVDLSHFLKEKGLGSLHSQRQDILDLDWIYHTQY 120  
|||||  
QY 230 FPDQNYTPGCVRYPLTFGWCYKLVPEPKVEANKGENTSLHPVSLHGMDDPEREV 289  
|||||  
DB 121 FPDQNYTPGCVRYPLTFGWCYKLVPEPKVEANKGENTSLHPVSLHGMDDPEREV 180  
|||||  
QY 290 LEWRDLSRLAFHHVARELHPEYFKNC 315  
|||||  
DB 181 LEWRDLSRLAFHHVARELHPEYFKNC 206  
|||||  
RESULT 14  
Q89561 PRELIMINARY; PRT; 206 AA.  
ID Q89561;  
AC Q89561;  
DT 01-NOV-1996 (TremBLrel. 01, Created)  
DT 01-NOV-1996 (TremBLrel. 01, Last sequence update)  
DT 01-DEC-2001 (TremBLrel. 19, Last annotation update)  
DE NEGATIVE FACTOR (F-PROTEIN) (27 KDA PROTEIN).  
GN NEF.

OS Human immunodeficiency virus type 1.  
OC Viruses; Retroviral viruses; Retroviridae; Lentivirus.  
OX NCBI\_TaxID=11676;  
RN [1]  
RP SEQUENCE FROM N.A.  
RA Arens M.O., Ratner L., Joseph T., Bandres J.,  
RL Submitted (JAN-1996) to the EMBL/GenBank/DBJ databases.  
CC -I- SIMILARITY: NEF HAS GTPASE, GTP-BINDING AND AUTOPHOSPHORYLATING  
ACTIVITIES, IT SEEMS TO DOWN-REGULATE THE CD4(T4) ANTIGEN.  
DR EMBL; U44448; AAB38200.1; -;  
DR EMBL; U44447; AAB38199.1; -;  
DR HSP; Q70627; 2NEF.  
DR InterPro: IPR001558; F-protein.  
DR Pfam: PF00469; F-protein; 1.  
DR AIDS; GTP-binding; Myristate.  
KW AIDS; GTP-binding; Myristate.  
SQ SEQUENCE 206 AA; 23629 MW; BF467FB1B7147CDE CRC64;

Query Match 48.2%; Score 1087; DB 15; Length 206;  
Best Local Similarity 94.7%; Pred. No. 6.6e-80;  
Matches 195; Conservative 7; Mismatches 4; Indels 0; Gaps 0;  
QY 110 MGGKWSKSSVVGWPTVRMRRAEPAADGVGAASRDLEKKGSLHSQRQDILDLDWIYHTQY 169  
|||||  
DB 1 MGGKWSKSSVVGWPTVRMRRAEPAADGVGAASRDLEKKGSLHSQRQDILDLDWIYHTQY 60  
|||||  
QY 170 QEEVEGFPVTPQVPLRPMTYKAAVDLSHFLKEKGLGSLHSQRQDILDLDWIYHTQY 229  
|||||  
DB 61 QEEVEGFPVTPQVPLRPMTYKAAVDLSHFLKEKGLGSLHSQRQDILDLDWIYHTQY 120  
|||||  
QY 230 FPDQNYTPGCVRYPLTFGWCYKLVPEPKVEANKGENTSLHPVSLHGMDDPEREV 289  
|||||  
DB 121 FPDQNYTPGCVRYPLTFGWCYKLVPEPKVEANKGENTSLHPVSLHGMDDPEREV 180  
|||||  
QY 290 LEWRDLSRLAFHHVARELHPEYFKNC 315  
|||||  
DB 181 LEWRDLSRLAFHHVARELHPEYFKNC 206  
|||||

RESULT 15  
Q74913 PRELIMINARY; PRT; 206 AA.  
ID Q74913;  
AC Q74913;  
DT 01-NOV-1996 (TremBLrel. 01, Created)  
DT 01-NOV-1996 (TremBLrel. 01, Last sequence update)  
DT 01-DEC-2001 (TremBLrel. 19, Last annotation update)  
DE NEGATIVE FACTOR (F-PROTEIN) (27 KDA PROTEIN).  
GN NEF.  
OS Human immunodeficiency virus type 1.  
OC Viruses; Retroviral viruses; Retroviridae; Lentivirus.  
OX NCBI\_TaxID=11676;  
RN [1]  
RP SEQUENCE FROM N.A.  
RX MEDLINE=96400183; PubMed=8806559;  
RA Ratner L., Joseph T., Bandres J., Ghosh S., Heyden N.V., Templeton A.,  
Hahn B., Powderly W., Arens M.,  
RT "Sequence heterogeneity of Nef transcripts in HIV-1-infected subjects  
at different stages of disease."  
RL Virology 223:245-250(1996).  
CC -I- SIMILARITY: NEF HAS GTPASE, GTP-BINDING AND AUTOPHOSPHORYLATING  
ACTIVITIES, IT SEEMS TO DOWN-REGULATE THE CD4(T4) ANTIGEN.  
DR EMBL; U44453; AAB38205.1; -;  
DR HSP; P03406; 1EFN.  
DR InterPro: IPR001558; F-protein.  
DR Pfam: PF00469; F-protein; 1.  
DR AIDS; GTP-binding; Myristate.  
KW AIDS; GTP-binding; Myristate.  
SQ SEQUENCE 206 AA; 23649 MW; E5B97FA0B70FC175 CRC64;

Query Match 48.1%; Score 1085; DB 15; Length 206;  
Best Local Similarity 95.6%; Pred. No. 9.6e-80;  
Matches 196; Conservative 5; Mismatches 4; Indels 0; Gaps 0;

QY 110 MGGKWSKSSVVGWPTVRERMRRAEPAADGVAASRDLEKKGAISSNTAATNAACAWLEA 169  
DB 1 MGGKWSKSSVVGWPTVRERMRRAEPAADGVAASRDLEKKGAISSNTAATNAACAWLEA 60  
QY 170 QEEVEGFPVTPQVPLRPMTYKAADVLSHFLKEKGLGSLHSQRQDILDWIYHTQGY 229  
DB 61 QEEVEGFPVTPQVPLRPMTYKAADVLSHFLKEKGLGSLHSQRQDILDWIYHTQGY 120  
QY 230 FPDQNYTPGCVRYPLTFGWCYKLVPEPKVEANKGENTSLHPVSLHGMDDPEREV 289  
DB 121 FPDQNYTPGCVRYPLTFGWCYKLVPEPKVEANKGENTSLHPVSLHGMDDPEREV 180  
QY 290 LEWRFDSRLAFHHVARELHPEYFKN 314  
DB 181 LEWRFDSRLAFHHVARELHPEYFKN 205  
RESULT 16  
Q74917 PRELIMINARY; PRT; 206 AA.  
AC Q74917;  
DT 01-NOV-1996 (TReMBLrel. 01, Created)  
DT 01-NOV-1996 (TReMBLrel. 01, Last sequence update)  
DT 01-DEC-2001 (TReMBLrel. 19, Last annotation update)  
DE NEGATIVE FACTOR (F-PROTEIN) (27 KDA PROTEIN).  
GN NEF.  
OS Human immunodeficiency virus type 1.  
OC Viruses; Retrovirdae; Lentivirus.  
OX NCBI\_TaxID=11676;  
RN [1]  
RP SEQUENCE FROM N.A.  
RX MEDLINE=96400183; PubMed=8806559;  
RA Ratner L., Joseph T., Bandres J., Ghosh S., Heyden N.V., Templeton A.,  
Hahn B., Powderly W., Arens M.;  
RT "Sequence heterogeneity of Nef transcripts in HIV-1-infected subjects  
at different stages of disease."  
RL Virology 223:245-250(1996).  
CC -!- SIMILARITY: NEF HAS GTPASE, GTP-BINDING AND AUTOPHOSPHORYLATING  
CC ACTIVITIES. IT SEEMS TO DOWN-REGULATE THE CD4(T4) ANTIGEN.  
DR EMBL; U44450; AAB38210.1; -;  
DR HSSP; P03406; IEFN.  
DR InterPro; IPR001558; F-protein.  
DR Pfam; PF00469; F-protein; 1.  
KW AIDS; GTP-binding; Myristate.  
SQ SEQUENCE 206 AA; 23562 MW; 478BBFA675CBC203 CRC64;  
Query Match 47.9%; Score 1080; DB 15; Length 206;  
Best Local Similarity 94.7%; Pred. No. 2.4e-79;  
Matches 195; Conservative 5; Mismatches 6; Indels 0; Gaps 0;  
QY 110 MGGKWSKSSVVGWPTVRERMRRAEPAADGVAASRDLEKKGAISSNTAATNAACAWLEA 169  
DB 1 MGGKWSKSSVVGWPTVRERMRRAEPAADGVAASRDLEKKGAISSNTAATNAACAWLEA 60  
QY 170 QEEVEGFPVTPQVPLRPMTYKAADVLSHFLKEKGLGSLHSQRQDILDWIYHTQGY 229  
DB 61 QEEVEGFPVTPQVPLRPMTYKAADVLSHFLKEKGLGSLHSQRQDILDWIYHTQGY 120  
QY 230 FPDQNYTPGCVRYPLTFGWCYKLVPEPKVEANKGENTSLHPVSLHGMDDPEREV 289  
DB 121 FPDQNYTPGCVRYPLTFGWCYKLVPEPKVEANKGENTSLHPVSLHGMDDPEREV 180  
QY 290 LEWRFDSRLAFHHVARELHPEYFKN 315  
DB 181 LEWRFDSRLAFHHVARELHPEYFKN 206  
RESULT 17  
Q74914 PRELIMINARY; PRT; 206 AA.  
AC Q74914;  
DT 01-NOV-1996 (TReMBLrel. 01, Created)

DT 01-NOV-1996 (TReMBLrel. 01, Last sequence update)  
DT 01-DEC-2001 (TReMBLrel. 19, Last annotation update)  
DE NEGATIVE FACTOR (F-PROTEIN) (27 KDA PROTEIN).  
GN NEF.  
OS Human immunodeficiency virus type 1.  
OC Viruses; Retrovirdae; Lentivirus.  
OX NCBI\_TaxID=11676;  
RN [1]  
RP SEQUENCE FROM N.A.  
RX MEDLINE=96400183; PubMed=8806559;  
RA Ratner L., Joseph T., Bandres J., Ghosh S., Heyden N.V., Templeton A.,  
Hahn B., Powderly W., Arens M.;  
RT "Sequence heterogeneity of Nef transcripts in HIV-1-infected subjects  
at different stages of disease."  
RL Virology 223:245-250(1996).  
CC -!- SIMILARITY: NEF HAS GTPASE, GTP-BINDING AND AUTOPHOSPHORYLATING  
CC ACTIVITIES. IT SEEMS TO DOWN-REGULATE THE CD4(T4) ANTIGEN.  
DR EMBL; U44454; AAB38206.1; -;  
DR HSSP; P03406; IEFN.  
DR InterPro; IPR001558; F-protein.  
DR Pfam; PF00469; F-protein; 1.  
KW AIDS; GTP-binding; Myristate.  
SQ SEQUENCE 206 AA; 23615 MW; 08D97FA0B70FC17D CRC64;  
Query Match 47.8%; Score 1079; DB 15; Length 206;  
Best Local Similarity 95.1%; Pred. No. 2.9e-79;  
Matches 195; Conservative 5; Mismatches 5; Indels 0; Gaps 0;  
QY 110 MGGKWSKSSVVGWPTVRERMRRAEPAADGVAASRDLEKKGAISSNTAATNAACAWLEA 169  
DB 1 MGGKWSKSSVVGWPTVRERMRRAEPAADGVAASRDLEKKGAISSNTAATNAACAWLEA 60  
QY 170 QEEVEGFPVTPQVPLRPMTYKAADVLSHFLKEKGLGSLHSQRQDILDWIYHTQGY 229  
DB 61 QEEVEGFPVTPQVPLRPMTYKAADVLSHFLKEKGLGSLHSQRQDILDWIYHTQGY 120  
QY 230 FPDQNYTPGCVRYPLTFGWCYKLVPEPKVEANKGENTSLHPVSLHGMDDPEREV 289  
DB 121 FPDQNYTPGCVRYPLTFGWCYKLVPEPKVEANKGENTSLHPVSLHGMDDPEREV 180  
QY 290 LEWRFDSRLAFHHVARELHPEYFKN 314  
DB 181 LEWRFDSRLAFHHVARELHPEYFKN 205  
RESULT 18  
Q74915 PRELIMINARY; PRT; 206 AA.  
AC Q74915;  
DT 01-NOV-1996 (TReMBLrel. 01, Created)  
DT 01-NOV-1996 (TReMBLrel. 01, Last sequence update)  
DT 01-DEC-2001 (TReMBLrel. 19, Last annotation update)  
DE NEGATIVE FACTOR (F-PROTEIN) (27 KDA PROTEIN).  
GN NEF.  
OS Human immunodeficiency virus type 1.  
OC Viruses; Retrovirdae; Lentivirus.  
OX NCBI\_TaxID=11676;  
RN [1]  
RP SEQUENCE FROM N.A.  
RX MEDLINE=96400183; PubMed=8806559;  
RA Ratner L., Joseph T., Bandres J., Ghosh S., Heyden N.V., Templeton A.,  
Hahn B., Powderly W., Arens M.;  
RT "Sequence heterogeneity of Nef transcripts in HIV-1-infected subjects  
at different stages of disease."  
RL Virology 223:245-250(1996).  
CC -!- SIMILARITY: NEF HAS GTPASE, GTP-BINDING AND AUTOPHOSPHORYLATING  
CC ACTIVITIES. IT SEEMS TO DOWN-REGULATE THE CD4(T4) ANTIGEN.  
DR EMBL; U44455; AAB38207.1; -;  
DR HSSP; P03406; IEFN.  
DR InterPro; IPR001558; F-protein.  
DR Pfam; PF00469; F-protein; 1.  
KW AIDS; GTP-binding; Myristate.

SQ SEQUENCE 206 AA; 23648 MW; E5BD6FA0B70FC175 CRC64;

Query Match 47.8%; Score 1079; DB 15; Length 206;

Best Local Similarity 95.6%; Pred. No. 2.9e-79;

Matches 195; Conservative 5; Mismatches 4; Indels 0; Gaps 0;

Qy 110 MGGKWSKSSVGVPTVVRMRRAEPAADGVGAASRDLEKKGAISSNTAATNAACAWLEA 169

Db 1 MGGKWSKSSVGVPTVVRMRRAEPAADGVGAASRDLEKKGAISSNTAATNAACAWLEA 60

Qy 170 QEEVEGVFPVTPQVPLRPMTYKAAVDLSHFLKEKGLGSLHSORQDILDWIYHTQY 229

Db 61 QEEVEGVFPVTPQVPLRPMTYKAAVDLSHFLKEKGLGSLHSORQDILDWIYHTQY 120

Qy 230 FPDQNYTPGPGVRYPLTFGWCYKLVPEPKVEEANKGENTSLHPVSLHGMDDPEREV 289

Db 121 FPDQNYTPGPGVRYPLTFGWCYKLVPEPKVEEANKGENTSLHPVSLHGMDDPEREV 180

Qy 290 LEWRDLSRLAFHHVARELHPEYK 313

Db 181 LEWRDLSRLAFHHVARELHPEYK 204

RESULT 19

Q74909

ID Q74909 PRELIMINARY; PRT; 206 AA.

AC Q74909;

DT 01-NOV-1996 (Tremblrel. 01, Created)

DT 01-NOV-1996 (Tremblrel. 01, Last sequence update)

DT 01-DEC-2001 (Tremblrel. 19, Last annotation update)

DE NEGATIVE FACTOR (F-PROTEIN) (27 KDA PROTEIN).

GN NEF.

OS Human immunodeficiency virus type 1.

OC Viruses; Retroviral viruses; Retroviridae; Lentivirus.

OX NCBI\_TaxID=11676;

RN [1]

RP SEQUENCE FROM N.A.

RX MEDLINE=96400183; PubMed=8806559;

RA Ratner L., Joseph T., Bandres J., Ghosh S., Heyden N.V., Templeton A.,

RA Hahn B., Powderly W., Arens M.;

RT "Sequence heterogeneity of Nef transcripts in HIV-1-infected subjects

RT at different stages of disease.";

RL Virology 223:245-250(1996).

CC -1- SIMILARITY: NEF HAS GTPASE, GTP-BINDING AND AUTOPHOSPHORYLATING

CC ACTIVITIES, IT SEEMS TO DOWN-REGULATE THE CD4 (T4) ANTIGEN.

DR EMBL; U44449; AAB38201.1; -

DR HSSP; Q70627; 2NEF.

DR InterPro; IPR001558; F-protein.

DR Pfam; PF00469; F-protein; 1.

KW AIDS: GTP-binding; Myristate.

SQ SEQUENCE 206 AA; 23488 MW; F0596D6FAAA81A05 CRC64;

Query Match 47.8%; Score 1077; DB 15; Length 206;

Best Local Similarity 94.2%; Pred. No. 4.2e-79;

Matches 194; Conservative 7; Mismatches 5; Indels 0; Gaps 0;

Qy 110 MGGKWSKSSVGVPTVVRMRRAEPAADGVGAASRDLEKKGAISSNTAATNAACAWLEA 169

Db 1 MGGKWSKSSVGVPTVVRMRRAEPAADGVGAASRDLEKKGAISSNTAATNAACAWLEA 60

Qy 170 QEEVEGVFPVTPQVPLRPMTYKAAVDLSHFLKEKGLGSLHSORQDILDWIYHTQY 229

Db 61 QEEVEGVFPVTPQVPLRPMTYKAAVDLSHFLKEKGLGSLHSORQDILDWIYHTQY 120

Qy 230 FPDQNYTPGPGVRYPLTFGWCYKLVPEPKVEEANKGENTSLHPVSLHGMDDPEREV 289

Db 121 FPDQNYTPGPGVRYPLTFGWCYKLVPEPKVEEANKGENTSLHPVSLHGMDDPEREV 180

Qy 290 LEWRDLSRLAFHHVARELHPEYK 315

Db 181 LEWRDLSRLAFHHVARELHPEYK 206

RESULT 20

Q9QPN3

ID Q9QPN3 PRELIMINARY; PRT; 202 AA.

AC Q9QPN3;

DT 01-MAY-2000 (Tremblrel. 13, Created)

DT 01-MAY-2000 (Tremblrel. 13, Last sequence update)

DT 01-DEC-2001 (Tremblrel. 19, Last annotation update)

DE NEGATIVE FACTOR (F-PROTEIN) (27 KDA PROTEIN) (FRAGMENT).

GN NEF.

OS Human immunodeficiency virus type 1.

OC Viruses; Retroviral viruses; Retroviridae; Lentivirus.

OX NCBI\_TaxID=11676;

RN [1]

RP SEQUENCE FROM N.A.

RC STRAIN=HIV-1LAI;

RA Piedade J., Esteves A., Parreira R., Venenno T., Barros M.F.,

RA Canas-Ferreira W.F.;

RT "Cloning and expression of HIV-1 nef gene in the carrier-adjutant

RT pVb3 expression system based on the major lipoprotein (Opr1) from the

RT outer membrane of Pseudomonas aeruginosa.";

RL Submitted (JUL-1999) to the EMBL/GenBank/DBJ databases.

CC -1- SIMILARITY: NEF HAS GTPASE, GTP-BINDING AND AUTOPHOSPHORYLATING

CC ACTIVITIES, IT SEEMS TO DOWN-REGULATE THE CD4 (T4) ANTIGEN.

DR EMBL; AF166101; AAD47831.1; -

DR HSSP; P03406; LEFN.

DR InterPro; IPR001558; F-protein.

DR Pfam; PF00469; F-protein; 1.

KW AIDS: GTP-binding; Myristate.

FT NON\_TER 202

SQ SEQUENCE 202 AA; 23033 MW; CD61DFA6F386CC89 CRC64;

Query Match 47.7%; Score 1076; DB 15; Length 202;

Best Local Similarity 97.5%; Pred. No. 5e-79;

Matches 197; Conservative 0; Mismatches 5; Indels 0; Gaps 0;

Qy 110 MGGKWSKSSVGVPTVVRMRRAEPAADGVGAASRDLEKKGAISSNTAATNAACAWLEA 169

Db 1 MGGKWSKSSVGVPTVVRMRRAEPAADGVGAASRDLEKKGAISSNTAATNAACAWLEA 60

Qy 170 QEEVEGVFPVTPQVPLRPMTYKAAVDLSHFLKEKGLGSLHSORQDILDWIYHTQY 229

Db 61 QEEVEGVFPVTPQVPLRPMTYKAAVDLSHFLKEKGLGSLHSORQDILDWIYHTQY 120

Qy 230 FPDQNYTPGPGVRYPLTFGWCYKLVPEPKVEEANKGENTSLHPVSLHGMDDPEREV 289

Db 121 FPDQNYTPGPGVRYPLTFGWCYKLVPEPKVEEANKGENTSLHPVSLHGMDDPEREV 180

Qy 290 LEWRDLSRLAFHHVARELHPEY 311

Db 181 LEWRDLSRLAFHHVARELHPEY 202

RESULT 21

Q9Q596

ID Q9Q596 PRELIMINARY; PRT; 206 AA.

AC Q9Q596;

DT 01-MAY-2000 (Tremblrel. 13, Created)

DT 01-MAY-2000 (Tremblrel. 13, Last sequence update)

DT 01-DEC-2001 (Tremblrel. 19, Last annotation update)

DE NEGATIVE FACTOR (F-PROTEIN) (27 KDA PROTEIN).

GN NEF.

OS Human immunodeficiency virus type 1.

OC Viruses; Retroviral viruses; Retroviridae; Lentivirus.

OX NCBI\_TaxID=11676;

RN [1]

RP SEQUENCE FROM N.A.

RA Wang B., Saksena N.K.;

RT "HIV-1 Strains from a cohort of American subjects reveal the presence

RT of a V2 region extension unique to slow progressors and non-

RT progressors.";



RP SEQUENCE FROM N.A.  
 RC STRAIN-PATIENT 27;  
 RA McPhree D.A., Greenway A.L., Holloway G., Smith K., Deacon N.,  
 RA Pemberton L., Brew B.J.;  
 RT "Anomalies in Nef expression within the central nervous system of HIV-  
 RT 1 positive individuals/AIDS patients with or without AIDS dementia  
 RT complex";  
 RL J. Neurovirol. 4:0-0(1998).  
 CC -I- SIMILARITY: NEF HAS GTPASE, GTP-BINDING AND AUTOPHOSPHORYLATING  
 CC ACTIVITIES, IT SEEMS TO DOWN-REGULATE THE CD4(T4) ANTIGEN.  
 DR EMBL; AF064676; AAC18377.1; -  
 DR HSSP; P03406; 1EPN.  
 DR InterPro; IPR001558; F-protein.  
 DR Pfam; PF00469; F-protein; 1.  
 KW AIDS; GTP-binding; Myristate.  
 SQ SEQUENCE 206 AA; 23499 MW; DEE21CAF05891D6B CRC64;

Query Match 47.4%; Score 1068; DB 15; Length 206;  
 Best Local Similarity 91.7%; Pred. No. 2.2e-78;  
 Matches 189; Conservative 14; Mismatches 3; Indels 0; Gaps 0;  
 Qy 110 MGGKWSKSSVVGWPTVRMRRAEPAADGVGAASRDLEKHAITSSNTAATNAACAWLEA 169  
 Db 1 MGGKWSKSSVVGWPTVRMRRAEPAADGVGAASRDLEKHAITSSNTAATNAACAWLEA 60  
 Qy 170 QEEVEGFPVTPQVPLRPMTYKAAVDLSHFLKEKGGLEGLHSORRODILDWIYHTQGY 229  
 Db 61 QKEEEVGFPVTPQVPLRPMTYKAAVDLSHFLKEKGGLEGLHSORRODILDWIYHTQGY 120  
 Qy 230 FPDQNTYTPGVRYPPLTFGWCYKLVPEPKVEEANKGENTSLHPVSLHGMDDPEREV 289  
 Db 121 FPDQNTYTPGVRYPPLTFGWCYKLVPEPKVEEANKGENTSLHPVSLHGMDDPEREV 180  
 Qy 290 LEWRFDLSLAFHHVARELHPEYKNC 315  
 Db 181 LEWRFDLSLAFHHVARELHPEYKDC 206

RESULT 25  
 QDQOI ID Q9DQOI PRELIMINARY; PRT; 206 AA.  
 AC Q9DQOI;  
 DT 01-MAR-2001 (TRENBLrel. 16, Created)  
 DT 01-MAR-2001 (TRENBLrel. 16, Last sequence update)  
 DT 01-DEC-2001 (TRENBLrel. 19, Last annotation update)  
 DE NEGATIVE FACTOR (F-PROTEIN) (27 KDA PROTEIN).  
 GN NEF.  
 OS Human immunodeficiency virus type 1.  
 OC Viruses; Retrovirdae; Retroviridae; Lentivirus.  
 OX NCBI\_TaxID=11676;  
 RN [1]  
 RC SEQUENCE FROM N.A.  
 RC STRAIN=03003;  
 RX MEDLINE-21002575; PubMed-11118071;  
 RA Geffin R., Wolf D., Muller R., Hill M.D., Stellweg E., Freitag M.,  
 RA Sass G., Scott G.B., Baur A.S.;  
 RT "Functional and structural defects in HIV-1 nef genes derived from  
 RT pediatric long-term survivors";  
 RL AIDS Res. Hum. Retroviruses 16:1855-1869(2000).  
 CC -I- SIMILARITY: NEF HAS GTPASE, GTP-BINDING AND AUTOPHOSPHORYLATING  
 CC ACTIVITIES, IT SEEMS TO DOWN-REGULATE THE CD4(T4) ANTIGEN.  
 DR EMBL; AF252911; AAG34580.1; -  
 DR HSSP; P03406; 1EPN.  
 DR InterPro; IPR001558; F-protein.  
 DR Pfam; PF00469; F-protein; 1.  
 KW AIDS; GTP-binding; Myristate.  
 SQ SEQUENCE 206 AA; 23464 MW; B72EOAA736374B64 CRC64;

Query Match 47.2%; Score 1065; DB 15; Length 206;  
 Best Local Similarity 93.2%; Pred. No. 3.9e-78;  
 Matches 192; Conservative 9; Mismatches 5; Indels 0; Gaps 0;

Qy 110 MGGKWSKSSVVGWPTVRMRRAEPAADGVGAASRDLEKHAITSSNTAATNAACAWLEA 169  
 Db 1 MGGKWSKSSVVGWPTVRMRRAEPAADGVGAASRDLEKHAITSSNTAATNAACAWLEA 60  
 Qy 170 QEEVEGFPVTPQVPLRPMTYKAAVDLSHFLKEKGGLEGLHSORRODILDWIYHTQGY 229  
 Db 61 QEEVEGFPVTPQVPLRPMTYKAAVDLSHFLKEKGGLEGLHSORRODILDWIYHTQGY 120  
 Qy 230 FPDQNTYTPGVRYPPLTFGWCYKLVPEPKVEEANKGENTSLHPVSLHGMDDPEREV 289  
 Db 121 FPDQNTYTPGVRYPPLTFGWCYKLVPEPKVEEANKGENTSLHPVSLHGMDDPEREV 180  
 Qy 290 LEWRFDLSLAFHHVARELHPEYKNC 315  
 Db 181 LEWRFDLSLAFHHVARELHPEYKDC 206

RESULT 26  
 Q9W7UO ID Q9W7UO PRELIMINARY; PRT; 206 AA.  
 AC Q9W7UO;  
 DT 01-NOV-1999 (TRENBLrel. 12, Created)  
 DT 01-NOV-1999 (TRENBLrel. 12, Last sequence update)  
 DT 01-DEC-2001 (TRENBLrel. 19, Last annotation update)  
 DE NEGATIVE FACTOR (F-PROTEIN) (27 KDA PROTEIN).  
 GN NEF.  
 OS Human immunodeficiency virus type 1.  
 OC Viruses; Retrovirdae; Retroviridae; Lentivirus.  
 OX NCBI\_TaxID=11676;  
 RN [1]  
 RC SEQUENCE FROM N.A.  
 RC STRAIN-PATIENT 27;  
 RA McPhree D.A., Greenway A.L., Holloway G., Smith K., Deacon N.,  
 RA Pemberton L., Brew B.J.;  
 RT "Anomalies in Nef expression within the central nervous system of HIV-  
 RT 1 positive individuals/AIDS patients with or without AIDS dementia  
 RT complex";  
 RL J. Neurovirol. 4:0-0(1998).  
 CC -I- SIMILARITY: NEF HAS GTPASE, GTP-BINDING AND AUTOPHOSPHORYLATING  
 CC ACTIVITIES, IT SEEMS TO DOWN-REGULATE THE CD4(T4) ANTIGEN.  
 DR EMBL; AF064675; AAC18376.1; -  
 DR EMBL; AF064674; AAC18375.1; -  
 DR HSSP; P03406; 1EPN.  
 DR InterPro; IPR001558; F-protein.  
 DR Pfam; PF00469; F-protein; 1.  
 KW AIDS; GTP-binding; Myristate.  
 SQ SEQUENCE 206 AA; 23514 MW; DC681CAF05891D6B CRC64;

Query Match 47.2%; Score 1064; DB 15; Length 206;  
 Best Local Similarity 91.7%; Pred. No. 4.7e-78;  
 Matches 189; Conservative 13; Mismatches 4; Indels 0; Gaps 0;

Qy 110 MGGKWSKSSVVGWPTVRMRRAEPAADGVGAASRDLEKHAITSSNTAATNAACAWLEA 169  
 Db 1 MGGKWSKSSVVGWPTVRMRRAEPAADGVGAASRDLEKHAITSSNTAATNAACAWLEA 60  
 Qy 170 QEEVEGFPVTPQVPLRPMTYKAAVDLSHFLKEKGGLEGLHSORRODILDWIYHTQGY 229  
 Db 61 QKEEEVGFPVTPQVPLRPMTYKAAVDLSHFLKEKGGLEGLHSORRODILDWIYHTQGY 120  
 Qy 230 FPDQNTYTPGVRYPPLTFGWCYKLVPEPKVEEANKGENTSLHPVSLHGMDDPEREV 289  
 Db 121 FPDQNTYTPGVRYPPLTFGWCYKLVPEPKVEEANKGENTSLHPVSLHGMDDPEREV 180  
 Qy 290 LEWRFDLSLAFHHVARELHPEYKNC 315  
 Db 181 LEWRFDLSLAFHHVARELHPEYKDC 206

RESULT 27  
 Q93010

```
ID O93010 PRELIMINARY; PRT: 206 AA.
AC O93010;
DT 01-NOV-1998 (Tremblrel. 08, Created)
DT 01-NOV-1998 (Tremblrel. 08, Last sequence update)
DT 01-DEC-2001 (Tremblrel. 19, Last annotation update)
DE NEGATIVE FACTOR (F-PROTEIN) (27 KDA PROTEIN).
GN NEF.
OS Human immunodeficiency virus type 1.
OC Viruses; Retrovirdae; Retroviridae; Lentivirus.
OX NCBI_TaxID=11676;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN-PATIENT 9;
RX MEDLINE=98097260; PubMed=9436760;
RA Kang M.R., Cho Y.K., Chun J., Kim Y.B., Lee I., Lee H.J., Kim S.H.,
RA Kim Y.K., Yoon K., Yang J.M., Kim J.M., Shin Y.O., Kang C., Lee J.S.,
RA Choi K.W., Kim D.G., Fitch W.M., Kim S.;
RT "Phylogenetic analysis of the nef gene reveals a distinctive
RT monophyletic clade in Korean HIV-1 cases."
RL J. Acquir. Immune Defic. Syndr. Hum. Retrovirol. 17:58-68(1998).
CC -1- SIMILARITY: NEF HAS GTPASE, GTP-BINDING AND AUTOPHOSPHORYLATING
CC ACTIVITIES, IT SEEMS TO DOWN-REGULATE THE CD4(T4) ANTIGEN.
DR EMBL; AF063922; AAC17893.1; -.
DR HSSP; P03406; IEFN.
DR InterPro; IPR001558; F-protein.
DR Pfam; PF00469; F-protein; 1.
DR AIDS; GTP-binding; Myristate.
KW SEQUENCE 206 AA; 23348 MW; 4D3AEDF55FAE93E3 CRC64;
SQ
```

```
Query Match 47.1%; Score 1062; DB 15; Length 206;
Best Local Similarity 94.6%; Pred. No. 6.8e-78;
Matches 194; Conservative 3; Mismatches 8; Indels 0; Gaps 0;

QY 110 MGKSKSSVVGWPTVRRMRRAEPAADGVGAASRDLEKHGAISSNTAATNAACAWLEA 169
DB 1 MGKSKSSVIGHPVRRMRRAEPAADGVGAASRDLEKHGAISSNTAATNAACAWLEA 60

QY 170 QEEVEGFPVTPQVLRPMPTKYAAVDLSHFLKKGLEGLIHSQRQDILDWIYHTQGY 229
DB 61 QEEVEGFPVTPQVLRPMPTKYAAVDLSHFLKKGLEGLIHSQRQDILDWIYHTQGY 120

QY 230 FPDQNYTPGQVRYPLTFGWCYKLVPEPKVEEANKGENTSLHPVSLHGMDDEPREV 289
DB 121 FPDQNYTPGQVRYPLTFGWCYKLVPEPKVEEANKGENTSLHPVSLHGMDDEPREV 180

QY 290 LEWRFDLSRLAFHHVARELHPEYKNC 314
DB 181 LEWRFDLSRLAFHHVARELHPEYKNC 205

RESULT 28
O93012 PRELIMINARY; PRT: 206 AA.
AC O93012;
DT 01-NOV-1998 (Tremblrel. 08, Created)
DT 01-NOV-1998 (Tremblrel. 08, Last sequence update)
DT 01-DEC-2001 (Tremblrel. 19, Last annotation update)
DE NEGATIVE FACTOR (F-PROTEIN) (27 KDA PROTEIN).
GN NEF.
OS Human immunodeficiency virus type 1.
OC Viruses; Retrovirdae; Retroviridae; Lentivirus.
OX NCBI_TaxID=11676;
RN [1]
RP SEQUENCE FROM N.A.
RA McPhée D.A., Greenway A.L., Holloway G., Smith K., Deacon N.,
RA Pemberton L., Brew B.J.;
RT "Anomalies in Nef expression within the central nervous system of HIV-
RT 1 positive individuals/AIDS patients with or without AIDS dementia
RT complex."
RL J. Neurovirol. 4:0-0(1998).
CC -1- SIMILARITY: NEF HAS GTPASE, GTP-BINDING AND AUTOPHOSPHORYLATING
```

```
CC ACTIVITIES, IT SEEMS TO DOWN-REGULATE THE CD4(T4) ANTIGEN.
DR EMBL; AF064673; AAC18374.1; -.
DR HSSP; P03406; IEFN.
DR InterPro; IPR001558; F-protein.
DR Pfam; PF00469; F-protein; 1.
DR AIDS; GTP-binding; Myristate.
KW SEQUENCE 206 AA; 23472 MW; 2255A447ECE85456 CRC64;
SQ
```

```
Query Match 47.1%; Score 1061; DB 15; Length 206;
Best Local Similarity 91.7%; Pred. No. 8.2e-78;
Matches 189; Conservative 12; Mismatches 5; Indels 0; Gaps 0;

QY 110 MGKSKSSVVGWPTVRRMRRAEPAADGVGAASRDLEKHGAISSNTAATNAACAWLEA 169
DB 1 MGKSKSSVVGWPTVRRMRRAEPAADGVGAASRDLEKHGAISSNTAATNAACAWLEA 60

QY 170 QEEVEGFPVTPQVLRPMPTKYAAVDLSHFLKKGLEGLIHSQRQDILDWIYHTQGY 229
DB 61 QEEVEGFPVTPQVLRPMPTKYAAVDLSHFLKKGLEGLIHSQRQDILDWIYHTQGY 120

QY 230 FPDQNYTPGQVRYPLTFGWCYKLVPEPKVEEANKGENTSLHPVSLHGMDDEPREV 289
DB 121 FPDQNYTPGQVRYPLTFGWCYKLVPEPKVEEANKGENTSLHPVSLHGMDDEPREV 180

QY 290 LEWRFDLSRLAFHHVARELHPEYKNC 315
DB 181 LEWRFDLSRLAFHHVARELHPEYKNC 206

RESULT 29
Q9Q597 PRELIMINARY; PRT: 206 AA.
AC Q9Q597;
DT 01-MAY-2000 (Tremblrel. 13, Created)
DT 01-MAY-2000 (Tremblrel. 13, Last sequence update)
DT 01-DEC-2001 (Tremblrel. 19, Last annotation update)
DE NEGATIVE FACTOR (F-PROTEIN) (27 KDA PROTEIN).
GN NEF.
OS Human immunodeficiency virus type 1.
OC Viruses; Retrovirdae; Retroviridae; Lentivirus.
OX NCBI_TaxID=11676;
RN [1]
RP SEQUENCE FROM N.A.
RA Wang B., Saksena N.K.;
RT "HIV-1 strains from a cohort of American subjects reveal the presence
RT of a V2 region extension unique to slow progressors and non-
RT progressors."
RL AIDS 0:0-0(2000).
CC -1- SIMILARITY: NEF HAS GTPASE, GTP-BINDING AND AUTOPHOSPHORYLATING
CC ACTIVITIES, IT SEEMS TO DOWN-REGULATE THE CD4(T4) ANTIGEN.
DR EMBL; AF203196; AAF25318.1; -.
DR HSSP; P03406; IEFN.
DR InterPro; IPR001558; F-protein.
DR Pfam; PF00469; F-protein; 1.
DR AIDS; GTP-binding; Myristate.
KW SEQUENCE 206 AA; 23599 MW; 9C887A58541A9905 CRC64;
SQ
```

```
Query Match 47.0%; Score 1060; DB 15; Length 206;
Best Local Similarity 92.7%; Pred. No. 9.9e-78;
Matches 191; Conservative 7; Mismatches 8; Indels 0; Gaps 0;

QY 110 MGKSKSSVVGWPTVRRMRRAEPAADGVGAASRDLEKHGAISSNTAATNAACAWLEA 169
DB 1 MGKSKSRSGVWPTVRRMRRAEPAADGVGAASRDLEKHGAISSNTPTNNADCAWLEA 60

QY 170 QEEVEGFPVTPQVLRPMPTKYAAVDLSHFLKKGLEGLIHSQRQDILDWIYHTQGY 229
DB 61 QEEVEGFPVTPQVLRPMPTKYAAVDLSHFLKKGLEGLIHSQRQDILDWIYHTQGY 120

QY 230 FPDQNYTPGQVRYPLTFGWCYKLVPEPKVEEANKGENTSLHPVSLHGMDDEPREV 289
DB 121 FPDQNYTPGQVRYPLTFGWCYKLVPEPKVEEANKGENTSLHPVSLHGMDDEPREV 180
```

Db 121 FPDQNTYTPGIRYPLTFGWCYKLVPEQEKVKKANEKNTSLHPMSLHGMDDPREV 180  
Qy 290 LEWRFDSRLAFHHVARELHPEYFKNC 315  
Db 181 LEWRFDSRLAFHHVARELHPEYFKNC 206

## RESULT 30

Q9W7X2  
ID Q9W7X2 PRELIMINARY; PRT; 206 AA.  
AC Q9W7X2;  
DT 01-NOV-1999 (TREMBLrel. 12, Created)  
DT 01-NOV-1999 (TREMBLrel. 12, Last sequence update)  
DT 01-DEC-2001 (TREMBLrel. 19, Last annotation update)  
DE NEGATIVE FACTOR (F-PROTEIN) (27 KDA PROTEIN).  
GN NEF.  
OS Human immunodeficiency virus type 1.  
OC Viruses; Retrovirdae; Retroviridae; Lentivirus.  
OX NCBI\_TaxID=11676;  
RN (1)  
RP SEQUENCE FROM N.A.  
RC STRAIN-PATIENT 27;  
RA McPhee D.A., Greenway A.L., Holloway G., Smith K., Deacon N.,  
Pemberton L., Brew B.J.;  
RT "Anomalies in Nef expression within the central nervous system of HIV-  
1 positive individuals/AIDS patients with or without AIDS dementia  
complex".  
RL J. Neurovirol. 4:0-0(1998).  
CC -1- SIMILARITY: NEF HAS GTPASE, GTP-BINDING AND AUTOPHOSPHORYLATING  
ACTIVITIES, IT SEEMS TO DOWN-REGULATE THE CD4(T4) ANTIGEN.  
DR EMBL; AF064677; AAC18378.1; -.  
DR HSSP; P03406; 1EFN.  
DR InterPro: IPR001558; F-protein.  
DR Pfam; PF00469; F-protein; 1.  
KW AIDS; GTP-binding; Myristate;  
SQ SEQUENCE 206 AA; 23442 MW; DC6A96AF05891D6B CRC64;

Query Match 46.9%; Score 1057; DB 15; Length 206;  
Best Local Similarity 91.3%; Pred. No. 1.7e-77;  
Matches 188; Conservative 13; Mismatches 5; Indels 0; Gaps 0;  
Qy 110 MGGKWSKSSVVGWPTVRRMRRAEPADGVGAASRDLEKKGAITSSNTAATNAACAWLEA 169  
Db 1 MGGKWSKSSVVGWPTVRRMRRAEPADGVGAASRDLEKKGAITSSNTAATNAACAWLEA 60  
Qy 170 QEEVEGFPVTPQVPLRPMYKAAVDLSHFLKXGGLGLHSORRQDILDWIYHTGY 229  
Db 61 QKEEVGFPVRPQVPLRPMYKAAVDLSHFLKXGGLGLHSORRQDILDWIYHTGY 120  
Qy 230 FPDQNTYTPGIRYPLTFGWCYKLVPEQEKVKKANEKNTSLHPMSLHGMDDPREV 289  
Db 121 FPDQNTYTPGIRYPLTFGWCYKLVPEQEKVKKANEKNTSLHPMSLHGMDDPREV 180  
Qy 290 LEWRFDSRLAFHHVARELHPEYFKNC 315  
Db 181 LEWRFDSRLAFHHVARELHPEYFKNC 206

Search completed: August 26, 2002, 08:14:56  
Job time: 384 sec



GenCore version 4.5  
Copyright (c) 1993 - 2000 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: August 26, 2002, 08:12:27 ; Search time 99.16 Seconds  
(without alignments)  
462.621 Million cell updates/sec

Title: US-09-509-239-21

Perfect score: 2264

Sequence: 1 MDPSSSSNANTQMSDKI.....QSRGDPGPKETSGHHHHH 413

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 747574 seqs, 111073796 residues

Total number of hits satisfying chosen parameters: 747574

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : A\_Geneseq\_032802:\*\*

- 1: /SIDSI/gcgdata/hold-geneseq/geneseq-emb1/AA1980.DAT.\*
- 2: /SIDSI/gcgdata/hold-geneseq/geneseq-emb1/AA1981.DAT.\*
- 3: /SIDSI/gcgdata/hold-geneseq/geneseq-emb1/AA1982.DAT.\*
- 4: /SIDSI/gcgdata/hold-geneseq/geneseq-emb1/AA1983.DAT.\*
- 5: /SIDSI/gcgdata/hold-geneseq/geneseq-emb1/AA1984.DAT.\*
- 6: /SIDSI/gcgdata/hold-geneseq/geneseq-emb1/AA1985.DAT.\*
- 7: /SIDSI/gcgdata/hold-geneseq/geneseq-emb1/AA1986.DAT.\*
- 8: /SIDSI/gcgdata/hold-geneseq/geneseq-emb1/AA1987.DAT.\*
- 9: /SIDSI/gcgdata/hold-geneseq/geneseq-emb1/AA1988.DAT.\*
- 10: /SIDSI/gcgdata/hold-geneseq/geneseq-emb1/AA1989.DAT.\*
- 11: /SIDSI/gcgdata/hold-geneseq/geneseq-emb1/AA1990.DAT.\*
- 12: /SIDSI/gcgdata/hold-geneseq/geneseq-emb1/AA1991.DAT.\*
- 13: /SIDSI/gcgdata/hold-geneseq/geneseq-emb1/AA1992.DAT.\*
- 14: /SIDSI/gcgdata/hold-geneseq/geneseq-emb1/AA1993.DAT.\*
- 15: /SIDSI/gcgdata/hold-geneseq/geneseq-emb1/AA1994.DAT.\*
- 16: /SIDSI/gcgdata/hold-geneseq/geneseq-emb1/AA1995.DAT.\*
- 17: /SIDSI/gcgdata/hold-geneseq/geneseq-emb1/AA1996.DAT.\*
- 18: /SIDSI/gcgdata/hold-geneseq/geneseq-emb1/AA1997.DAT.\*
- 19: /SIDSI/gcgdata/hold-geneseq/geneseq-emb1/AA1998.DAT.\*
- 20: /SIDSI/gcgdata/hold-geneseq/geneseq-emb1/AA1999.DAT.\*
- 21: /SIDSI/gcgdata/hold-geneseq/geneseq-emb1/AA2000.DAT.\*
- 22: /SIDSI/gcgdata/hold-geneseq/geneseq-emb1/AA2001.DAT.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

#### SUMMARIES

Result No.	Score	Match	Length	ID	Description
1	2264	100.0	413	20 AAY02355	A representative L
2	2264	100.0	413	22 AAG63237	Amino acid sequenc
3	2246	99.2	411	22 AAG63235	Amino acid sequenc
4	2233	98.6	411	20 AAY02353	A representative L
5	1715.5	75.8	326	20 AAY02354	A representative L
6	1715.5	75.8	326	22 AAG63236	Amino acid sequenc
7	1697.5	75.0	324	20 AAY02352	A representative L
8	1697.5	75.0	324	22 AAG63234	Amino acid sequenc
9	1688	74.6	302	20 AAY02351	A representative H
10	1688	74.6	302	22 AAG63233	Amino acid sequenc
11	1675	74.0	302	20 AAY02357	A representative H

12	1675	74.0	302	22	AAG63239	Amino acid sequenc
13	1139.5	50.3	215	20	AAY02349	A representative H
14	1139.5	50.3	215	22	AAG63231	Amino acid sequenc
15	1120	49.5	206	21	AAY50795	Human NEF protein/
16	1116	49.3	206	14	AAR38893	Nef protein of HIV
17	1115	49.2	206	21	AAB10054	HIV-1 nef protein.
18	1114	49.1	206	20	AAR69326	HIV-1 nef protein
19	1111	49.1	206	7	AAP61515	Sequence of E' pro
20	1107	48.9	216	7	AAP60423	Sequence of LAV v1
21	1092	48.2	206	20	AAW90479	HTLV-III E' protei
22	1011	44.7	216	22	AAE04960	HIV-1 jrf1 nef pro
23	997	44.0	210	19	AAW53113	Protein 6 containe
24	997	44.0	210	21	AAW77299	HIV-1 (ATCC CRL 85
25	996	44.0	217	22	AAE04962	HIV-1 jrf1 Nef (G2
26	983	43.4	210	12	AAR12262	HIV-1 strain OYI o
27	978	43.2	237	22	AAE04961	Human tPA leader p
28	968	42.8	237	22	AAE04963	Human tPA leader p
29	952	42.0	206	21	AAB69363	HIV-1 non-subtype
30	948.5	41.9	3025	22	AAB86169	HIV-1 subtype C.pr
31	934	41.3	206	21	AAB69361	HIV-1 non-subtype
32	927.5	41.0	206	9	AAP81859	Sequence encoded b
33	910.5	40.2	207	21	AAB69359	HIV-1 non-subtype
34	903.5	39.9	207	11	AAR08407	Sequence deduced f
35	901.5	39.8	207	21	AAB69364	HIV-1 non-subtype
36	901.5	39.8	208	21	AAB69356	HIV-1 non-subtype
37	895.5	39.6	219	21	AAB69362	HIV-1 non-subtype
38	893.5	39.5	217	21	AAB69365	HIV-1 non-subtype
39	885.5	39.1	207	21	AAB69366	HIV-1 non-subtype
40	878.5	38.8	207	21	AAB69357	HIV-1 non-subtype
41	867	38.3	206	21	AAB69358	HIV-1 non-subtype
42	860	38.0	206	21	AAB69360	HIV-1 non-subtype
43	834.5	36.9	209	19	AAW72998	HIV isolate LAV.MA
44	831.5	36.7	209	9	AAP81866	Sequence encoded b
45	796	35.2	212	19	AAW68481	HIV-1 strain YBF30

#### ALIGNMENTS

RESULT 1

AAV02355

ID AAY02355 standard; Protein; 413 AA.

XX AC AAY02355;

XX DT 09-JUL-1999 (first entry)

XX DE A representative Lipod-Tat fusion protein.

XX KW HIV nef gene; fusion protein; HIV nef gene; Nef protein; Tat protein;

XX KW vaccine; HIV infection; protein D.

XX OS Synthetic.

XX OS Human immunodeficiency virus type 1.

XX PN WO9916884-AL.

XX PD 08-APR-1999.

XX PF 17-SEP-1998; 98WO-EP06040.

XX PR 26-SEP-1997; 97GB-0020585.

XX PR (SMIK ) SMITHKLINE BEECHAM BIOLOGICALS.

XX PI Bruck C, Godart SAG, Marchand M;

XX XX WPI; 1999-302282/25.

XX DR N-PSDB; AAX35691.

XX XX HIV Tat or Nef protein linked to a fusion partner

XX PS Disclosure; Fig 2; 66pp; English.

```

XX CC The present sequence represents a fusion protein comprising Lipod-HIV-1
CC CC Tat. The protein is exemplified the fusion proteins of
CC CC the invention. The specification also describes fusion proteins
CC CC comprising HIV-1 Nef protein. The fusion protein can be used in a
CC CC vaccine to prevent HIV infection.
XX CC
XX SQ Sequence 413 AA;

Query Match 100.0%; Score 2264; DB 20; Length 413;
Best Local Similarity 100.0%; Pred. No. 3.1e-205;
Matches 413; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 MDPSSHSNMANTQMSDKIIIAHRGASGYLPETHLESKALAFQAQADYLEQDLAMTKDG 60
Db 1 mdpsshsnmantqmsdkiiiahrgasgyipehtleskalafqaqadyleqdlamtkdg 60
Qy 61 RLVTIHDHFLDGLTDVAKKFPFHRKDGYYVIDFTLKEIOSLEMTENFETMGKWSKSS 120
Db 61 rlvtihdhflgltdvakkfphrkdgyyvifdtlkeioslmentenfetmgkwsks 120
Qy 121 VVGWPTVTRMRRAEPADGVAASRDLEKHAITSSNTAATNAACAWLEAQEEVEGFP 180
Db 121 vvgwptvtrmrreaepadgvaasrdlekhaitsntaatnaacawleaqeevegfp 180
Qy 181 VTPQVPLRPMTYKAAVDLSHFLKEKGLGLIHSORRODILDLWIYHTQGYPDMQNTYP 240
Db 181 vtpqvplrpmtykaavdlsfhflkekglglihsorrodildlwiyhtqgyfpmqnytp 240
Qy 241 GPGVRYPLTTFGCYKLVPEPDKVEANKGENTSLHPVSLHGMDDPEREVLEWRFDSRL 300
Db 241 gpgvryplttfgcylvpepdkveeankgentsllhpsvslhgmddperevlewrfdsrl 300
Qy 301 AFHHVARELHPEYFNKCTSEPVDPRLPEWKPFGSQPKTACTNCYCKKCCFHCQVCFITKA 360
Db 301 afhhvarelhpeyfnkctsevpdprlepwbkpgsqpkactncycckkccfchcvcfika 360
Qy 361 LGISYGRKKRRQRPRPGSQTHQVSLSKOPTSQSRGDPGTGPKETSGHHHHH 413
Db 361 lgisygrkrrqrprpgsqthqvskskoptsqsrqgdpdtpgkctsgghhhh 413

RESULT 2
AAG63237
ID AAG63237 standard; Protein: 413 AA.
XX AC
XX AC AAG63237;
XX DT 01-OCT-2001 (first entry)
XX DE Amino acid sequence of a His tagged ProTD-Nef-Tat fusion protein.
XX KW HIV; nef gene; tat gene; vaccine; gp120 gene; HIV viral load.
XX OS Synthetic.
XX OS Human immunodeficiency virus.
XX FH Key Location/Qualifiers
XX FT Peptide 1..111
XX FT /note= "ProTD fusion partner"
XX PN WO200154719-A2.
XX PD 02-AUG-2001.
XX PE 29-JAN-2001; 2001WO-EP00944.
XX PR 31-JAN-2000; 2000GB-0002200.
XX PR 14-APR-2000; 2000GB-0009336.
XX PR 06-JUN-2000; 2000GB-0013806.
XX PR 28-JUN-2000; 2000WO-EP05998.
XX OS

```

```

PA (SMIK ) SMITHKLINE BEECHAM BIOLOGICALS.
XX Voss G;
XX WPI: 2001-476172/51.
XX DR N-PSDB; AAH42881.
XX PT New use of a human immunodeficiency virus (HIV) Tat, or Nef, or Nef
XX linked to Tat (Nef-Tat) protein or polynucleotide and an HIV gp120
XX protein or polynucleotide for the manufacture of a vaccine -
XX Disclosure: Fig 1; 90pp; English.
XX CC The present sequence represents a His-tagged ProTD-Nef-Tat fusion
XX protein. The protein is expressed in Escherichia coli, and is used to
XX produce the vaccine of the invention. The specification describes
XX the use of HIV Tat, HIV Nef, or Nef-Tat; and HIV gp120 in the
XX manufacture of a vaccine. The vaccine is used for the prophylactic or
XX therapeutic immunization of humans against HIV. Tat, Nef or Nef-Tat act
XX in synergy with gp120 in the treatment and prevention of HIV. The
XX vaccine reduces the HIV viral load in HIV infected humans and results
XX in a maintenance of CD4+ levels over those levels found in the absence
XX of vaccination with HIV Tat, Nef or Nef-Tat and HIV gp120.
XX SQ Sequence 413 AA;

Query Match 100.0%; Score 2264; DB 22; Length 413;
Best Local Similarity 100.0%; Pred. No. 3.1e-205;
Matches 413; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 MDPSSHSNMANTQMSDKIIIAHRGASGYLPETHLESKALAFQAQADYLEQDLAMTKDG 60
Db 1 mdpsshsnmantqmsdkiiiahrgasgyipehtleskalafqaqadyleqdlamtkdg 60
Qy 61 RLVTIHDHFLDGLTDVAKKFPFHRKDGYYVIDFTLKEIOSLEMTENFETMGKWSKSS 120
Db 61 rlvtihdhflgltdvakkfphrkdgyyvifdtlkeioslmentenfetmgkwsks 120
Qy 121 VVGWPTVTRMRRAEPADGVAASRDLEKHAITSSNTAATNAACAWLEAQEEVEGFP 180
Db 121 vvgwptvtrmrreaepadgvaasrdlekhaitsntaatnaacawleaqeevegfp 180
Qy 181 VTPQVPLRPMTYKAAVDLSHFLKEKGLGLIHSORRODILDLWIYHTQGYPDMQNTYP 240
Db 181 vtpqvplrpmtykaavdlsfhflkekglglihsorrodildlwiyhtqgyfpmqnytp 240
Qy 241 GPGVRYPLTTFGCYKLVPEPDKVEANKGENTSLHPVSLHGMDDPEREVLEWRFDSRL 300
Db 241 gpgvryplttfgcylvpepdkveeankgentsllhpsvslhgmddperevlewrfdsrl 300
Qy 301 AFHHVARELHPEYFNKCTSEPVDPRLPEWKPFGSQPKTACTNCYCKKCCFHCQVCFITKA 360
Db 301 afhhvarelhpeyfnkctsevpdprlepwbkpgsqpkactncycckkccfchcvcfika 360
Qy 361 LGISYGRKKRRQRPRPGSQTHQVSLSKOPTSQSRGDPGTGPKETSGHHHHH 413
Db 361 lgisygrkrrqrprpgsqthqvskskoptsqsrqgdpdtpgkctsgghhhh 413

RESULT 3
AAG63235
ID AAG63235 standard; Protein: 411 AA.
XX AC
XX AC AAG63235;
XX DT 01-OCT-2001 (first entry)
XX DE Amino acid sequence of a His tagged LipD-Nef-Tat linked protein.
XX KW HIV; nef gene; tat gene; vaccine; gp120 gene; HIV viral load.
XX OS Synthetic.

```

OS Human immunodeficiency virus.

PH Key Location/Qualifiers  
FT Peptide 1..109  
TT /note= "ProtD fusion partner"

XX WO200154719-A2.

XX PD 02-AUG-2001.

XX PF 29-JAN-2001; 2001WO-EP00944.

XX PR 31-JAN-2000; 2000GB-0002200.

XX PR 14-APR-2000; 2000GB-0009336.

XX PR 06-JUN-2000; 2000GB-0013806.

XX PR 28-JUN-2000; 2000WO-EP05998.

XX PA (SMIK ) SMITHKLINE BEECHAM BIOLOGICALS.

XX PI Voss G;

XX WPI; 2001-476172/51.

XX DR N-PSDB; AAH42879.

XX New use of a human immunodeficiency virus (HIV) Tat, or Nef, or Nef

XX PT linked to Tat (Nef-Tat) protein or polynucleotide and an HIV gp120

XX PT protein or polynucleotide for the manufacture of a vaccine -

XX PS Disclosure: Fig 1; 90pp; English.

XX The present sequence represents a His-tagged Nef-Tat linked protein of  
CC HIV, with a lipidation signal sequence (lipod) which is removed after  
CC processing and a ProtD fusion partner. The protein is expressed in  
CC Escherichia coli, and is used to produce the vaccine of the invention.  
CC The specification describes the use of HIV Tat, HIV Nef, or Nef-Tat;  
CC and HIV gp120 in the manufacture of a vaccine. The vaccine is used for  
CC the prophylactic or therapeutic immunization of humans against HIV.  
CC Tat, Nef or Nef-Tat act in synergy with gp120 in the treatment and  
CC prevention of HIV. The vaccine reduces the HIV viral load in HIV  
CC infected humans and results in a maintenance of CD4+ levels over those  
CC levels found in the absence of vaccination with HIV Tat, Nef or Nef-Tat  
CC and HIV gp120.

XX Sequence 411 AA;

Query Match 99.2%; Score 2246; DB 22; Length 411;  
Best Local Similarity 100.0%; Pred. No. 1.5e-203;  
Matches 410; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 4 SHSSNMANTQMSDKIIIAHRCASGYLPEHTLESKALAFQAQADYLEQDLAMTKDRLV 63

DB 2 shssnmantqmsdkiiiahrgasgylpehtleskalfagqadyleqdlamtkdgrlv 61

QY 64 VTHDFLDGLTVAKKFPHRHRKDGRIYVIDFTLKEIQSLEMTENFETMGKWSKSSVVG 123

DB 62 vthdhfldgltdvakkfphrhrkdgryyvidftlkeiqslemtentfmgkwsksvvg 121

QY 124 WPTVRMRRAEPAADVGAASRDLEKHAITSNTAATNAACAWLEAEQEEVEGFPVTP 183

DB 122 wptvrermraepaadvgaasrdlekhgaitsntaataacawleaeqeevfpvtp 181

QY 184 QVPLRPMTYKAADVSHFLKEKGGLEGLTHSRRQDILDWYHTGYFPDQWNTYPPGG 243

DB 182 qvplrpmtykaadvshflkekggleglthsqrgldildwyhtgyfpdwgnytpgpg 241

QY 244 VRYPLTFGWCYKLVPEPKVEANKGENTSLHPVSLHGMDDPEREVLWFDSRLAFH 303

DB 242 vrypltfgcylkvppepkveankgentsllhpvslhgmddperewlwfdsrlafh 301

QY 304 HVARELHPYFKNCTSEPDVDRLEPKHFGSQPKTACTNCYCKKCCFHCQVCFITKALGI 363

DB 302 hvarelhpeyfnctsepdvdrlepkhfgsqpkactactncyckkccfhcqvcfitkalgi 361

QY 364 SYGRKKRRQRRRPPQSGOTHQVSLSKQPTSQSRGDPGTGPKETSGHHHHH 413  
DB 362 sygrkkrrqrrppqsgqthqvsksqptsqsgdtpgpketsghhhhh 411

#### RESULT 4

AAV02353

ID AAY02353 standard; Protein; 411 AA.

XX AC AAY02353;

XX DT 09-JUL-1999 (first entry)

XX DE A representative Lipod-Tat-His fusion protein.

XX KW HIV nef gene; fusion protein; HIV nef gene; Nef protein; Tat protein;

XX KW vaccine; HIV infection; protein D.

XX OS Synthetic.

OS Human immunodeficiency virus type 1.

XX PN WO9916884-A1.

XX PD 08-APR-1999.

XX PF 17-SEP-1998; 98WO-EP06040.

XX PR 26-SEP-1997; 97GB-0020585.

XX PA (SMIK ) SMITHKLINE BEECHAM BIOLOGICALS.

XX PI Bruck C, Godart SAG, Marchand M;

XX DR WPI; 1999-302282/25.

XX N-PSDB; AAX35689.

XX HIV Tat or Nef protein linked to a fusion partner

XX Disclosure: Fig 2; 66pp; English.

CC The present sequence represents a fusion protein comprising Lipod-Hiv-1  
CC Tat-His. The protein is exemplifies the fusion proteins of  
CC the invention. The specification also describes fusion proteins  
CC comprising HIV-1 Nef protein. The fusion protein can be used in a  
CC vaccine to prevent HIV infection.

XX Sequence 411 AA;

Query Match 98.6%; Score 2233; DB 20; Length 411;  
Best Local Similarity 99.8%; Pred. No. 2.6e-202;  
Matches 409; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 4 SHSSNMANTQMSDKIIIAHRCASGYLPEHTLESKALAFQAQADYLEQDLAMTKDRLV 63

DB 2 shssnmantqmsdkiiiahrgasgylpehtleskalfagqadyleqdlamtkdgrlv 61

QY 64 VTHDFLDGLTVAKKFPHRHRKDGRIYVIDFTLKEIQSLEMTENFETMGKWSKSSVVG 123

DB 62 vthdhfldgltdvakkfphrhrkdgryyvidftlkeiqslemtentfmgkwsksvvg 121

QY 124 WPTVRMRRAEPAADVGAASRDLEKHAITSNTAATNAACAWLEAEQEEVEGFPVTP 183

DB 122 wptvrermraepaadvgaasrdlekhgaitsntaataacawleaeqeevfpvtp 181

QY 184 QVPLRPMTYKAADVSHFLKEKGGLEGLTHSRRQDILDWYHTGYFPDQWNTYPPGG 243

DB 182 qvplrpmtykaadvshflkekggleglthsqrgldildwyhtgyfpdwgnytpgpg 241

QY 244 VRYPLTFGWCYKLVPEPKVEANKGENTSLHPVSLHGMDDPEREVLWFDSRLAFH 303

DB 242 vrypltfgcylkvppepkveankgentsllhpvslhgmddperewlwfdsrlafh 301

QY 304 HVARELHPEYFKNCTSEPVDRLEPDKVVEANKGENTSLHHPVSLHGMDDPEREVLHFRFDSRL 363  
 Db 302 hvarelhpeyfkntsetpdrlepwkhpqsgkptactncycckkccfchgcvcfittkalgi 361  
 QY 364 SYGRKKRRORRPPQGSQTHOVSLSKQPTQSOSRQDPTGPKETSGHHHHH 413  
 Db 362 sygrkrrrrppqgqthqvsksqptsgsrqdgtpgksetsgnhhhh 411

RESULT 5  
 AAY02354  
 ID AAY02354 standard; Protein: 326 AA.  
 XX AC AAY02354;  
 DT 09-JUL-1999 (first entry)  
 DE A representative Lipod-Nef fusion protein.  
 KW HIV nef gene; fusion protein; HIV nef gene; Nef protein; Tat protein;  
 KW vaccine; HIV infection; protein D.  
 XX XX  
 OS Synthetic.  
 OS Human immunodeficiency virus type 1.  
 XX WO9916884-A1.  
 PN 08-APR-1999.  
 PD 17-SEP-1998; 98WO-EP06040.  
 PF 26-SEP-1997; 97GB-0020585.  
 PR (SMIK ) SMITHKLINE BEECHAM BIOLOGICALS.  
 XX Bruck C, Godart SAG, Marchand M;  
 PI WPI: 1999-302282/25.  
 XX N-PSDB: AAX35690.  
 DR HIV Tat or Nef protein linked to a fusion partner  
 PT Disclosure; Fig 2; 66pp; English.  
 PS The present sequence represents a fusion protein comprising Lipod-HIV-1  
 XX Nef. The protein is exemplified by the fusion proteins of  
 CC the invention. The specification also describes fusion proteins  
 CC comprising HIV-1 Tat protein. The fusion protein can be used in a  
 CC vaccine to prevent HIV infection.  
 XX Sequence 326 AA;  
 SQ

Query Match 75.8%; Score 1715.5; DB 20; Length 326;  
 Best Local Similarity 78.9%; Pred. No. 1.5e-153;  
 Matches 326; Conservative 0; Mismatches 0; Indels 87; Gaps 1;

QY 1 MDPSHSSNMANTQMSKIIIAHAGSYLPETHLESKALAFQAQADYLEQDLAMTKDG 60  
 Db 1 mdpsshsnmantqmskiiiahgasylpethleskafafaqadyleqdlamtkdg 60  
 QY 61 RLUVTHDFDLGTLVAKKFPHRHRKDRYYVIDFTLKEIQSLEMTENFETMGKWSKSS 120  
 Db 61 rlviwhdfldgltvakkfphrhrkdrgyyvidftlkeiqslemtenfetmgkwsks 120  
 QY 121 VVGWPTVREMRRAEPADGVCASRDLEKHCALITSSNTAATNAACAWLEAQEEVEVGFP 180  
 Db 121 vvgwptvremrraepadgvcasrdlekhalitssntaataacawleaqeevevgfp 180  
 QY 181 VTPQVPLRPMTYKAAVDLSHFLKKEGLEGLIHSORRQDILDLWIYHTQGYPPDQWNTYP 240  
 Db 181 vtpqvpplrmtykaavdlsfhflkkekgleglihsqrrqdildlwyhtqgyfpdqwnytp 240

QY 241 GPGVRYPLTFTGWCYKLYVPEDPKVVEANKGENTSLHHPVSLHGMDDPEREVLHFRFDSRL 300  
 Db 241 gpgvryplftfgwcyklyvpedpkveeankgentsllhpsalhgmddperevlewrfsrl 300  
 QY 301 AFHVARELHPEYFKNCTSEPVDRLEPDKVVEANKGENTSLHHPVSLHGMDDPEREVLHFRFDSRL 360  
 Db 301 afhvarelhpeyfknc-----tsghhhhhh 317  
 QY 361 LGISYGRKKRRRRPPQGSQTHOVSLSKQPTQSOSRQDPTGPKETSGHHHHH 413  
 Db 318 -----tsghhhhhh 326

RESULT 6  
 AAG63236  
 ID AAG63236 standard; Protein: 326 AA.  
 XX AC AAG63236;  
 DT 01-OCT-2001 (first entry)  
 DE Amino acid sequence of a His tagged ProtD-Nef fusion protein.  
 KW HIV; nef gene; tat gene; vaccine; gp120 gene; HIV viral load.  
 XX XX  
 OS Synthetic.  
 OS Human immunodeficiency virus.  
 XX XX  
 FH Key Location/Qualifiers  
 FT Peptide 1..111  
 ET /note= "ProtD fusion partner"  
 XX WO200154719-A2.  
 PN 02-AUG-2001.  
 PD 29-JAN-2001; 2001WO-EP00944.  
 PF 31-JAN-2000; 2000GB-0002200.  
 PR 14-APR-2000; 2000GB-0009336.  
 PR 06-JUN-2000; 2000GB-0013806.  
 PR 28-JUN-2000; 2000WO-EP05998.  
 XX (SMIK ) SMITHKLINE BEECHAM BIOLOGICALS.  
 PA Voss G;  
 XX WPI: 2001-476172/51.  
 DR N-PSDB: AAH42880.  
 XX New use of a human immunodeficiency virus (HIV) Tat, or Nef, or Nef  
 PT linked to Tat (Nef-Tat) protein or polynucleotide and an HIV gp120  
 PT protein or polynucleotide for the manufacture of a vaccine -  
 PS Disclosure; Fig 1; 90pp; English.  
 XX The present sequence represents a His-tagged ProtD-Nef fusion protein.  
 CC The protein is expressed in Escherichia coli, and is used to  
 CC produce the vaccine of the invention. The specification describes  
 CC the use of HIV Tat, HIV Nef, or Nef-Tat; and HIV gp120 in the  
 CC manufacture of a vaccine. The vaccine is used for the prophylactic or  
 CC therapeutic immunization of humans against HIV. Tat, Nef or Nef-Tat act  
 CC in synergy with gp120 in the treatment and prevention of HIV. The  
 CC vaccine reduces the HIV viral load in HIV infected humans and results  
 CC in a maintenance of CD4+ levels over those levels found in the absence  
 CC of vaccination with HIV Tat, Nef or Nef-Tat and HIV gp120.  
 XX Sequence 326 AA;  
 SQ

Query Match 75.8%; Score 1715.5; DB 22; Length 326;  
 Best Local Similarity 78.9%; Pred. No. 1.5e-153;

Matches 326; Conservative 0; Mismatches 0; Indels 87; Gaps 1;

QY 1 MDPSSSHSNWANTQMSKIIIAHRCAGSYLPEHTLESKALAFQAQADYLEQDLAMTKDG 60  
 Db 1 mdpssshsnmantqmskdiiahrgasgylpehtleskalafagadyleqdlamtkdg 60

QY 61 RLVTVDHFLDGLTDAVAKFPFHRHRKDGYYVIDFTLKEIQSLEMTENFETMGKWSKSS 120  
 Db 61 rlvtvdhflldgltdavakffphrhrkdgryyvidftlkeiqslemtentfmggkwsks 120

QY 121 VVGWPTVTRMRRAEPAADGVGAASRDLEKKGAISSNTAATNAACAWLEAQEEVEEGFP 180  
 Db 121 vvgwptvtrmreraepaadvgaasrdlekgaitssntaataaacawleaqeeveegfp 180

QY 181 VTPQVPLRPMTYKAADVLSHFLKEKGLLEGLIHSORRQDILDILWIYHTGYPDPQNTYTP 240  
 Db 181 vtpqvplrpmtykaadvlsfhflkekgleglilhsgrrqddildilwiyhtgypdqnytp 240

QY 241 GPGVRYPTFTFGCYKLVPEPKVEEANKGENTSLHHPVSLHGMDDPEREVLWRFDSSL 300  
 Db 241 gpgvryptftfgwcyklvpvpdkveeankgentsllhpsvlhgmddperevlewrfdsl 300

QY 301 AFHHVARELHPEYFKNCTSEPVDPRLPEPMKHPGSPKTACTIONCYCKKCFHCQVCFITKA 360  
 Db 301 afhhvarelhpeyfknc-----tsghhhhhh 317

QY 361 LGISYGRKKRRRRRPPQGSQTHQVSLSKQPTSQSRGDPGPKETSGHHHHH 413  
 Db 318 -----tsghhhhhh 326

RESULT 7  
 AAY02352  
 ID AAY02352 standard; Protein; 324 AA.  
 AC AAY02352;  
 DT 09-JUL-1999 (first entry)  
 DE A representative Lipod-Nef-His fusion protein.  
 KW HIV nef gene; fusion protein; HIV nef gene; Nef protein; Tat protein;  
 KW vaccine; HIV infection; protein D.  
 OS Synthetic.  
 OS Human immunodeficiency virus type 1.  
 PN WO9916884-A1.  
 PD 08-APR-1999.  
 PF 17-SEP-1998; 98WO-EP06040.  
 PR 26-SEP-1997; 97GB-0020585.  
 XX (SMIK ) SMITHKLINE BEECHAM BIOLOGICALS.  
 PI Bruck C, Godart SAG, Marchand M;  
 DR WPI; 1999-302282/25.  
 DR N-PSDB; AAX35688.  
 PT HIV Tat or Nef protein linked to a fusion partner  
 XX Disclosure; Fig 2; 66pp; English.  
 CC The present sequence represents a fusion protein comprising Lipod-HIV-1  
 CC Nef-His. The protein is exemplified the fusion proteins of  
 CC the invention. The specification also describes fusion proteins  
 CC comprising HIV-1 Tat protein. The fusion protein can be used in a  
 CC vaccine to prevent HIV infection.  
 XX Sequence 324 AA;

Query Match 75.0%; Score 1697.5; DB 20; Length 324;  
 Best Local Similarity 78.8%; Pred. No. 7.5e-152;  
 Matches 323; Conservative 0; Mismatches 0; Indels 87; Gaps 1;

QY 4 SSHSNWANTQMSKIIIAHRCAGSYLPEHTLESKALAFQAQADYLEQDLAMTKDRLV 63  
 Db 2 sshsnmantqmskdiiahrgasgylpehtleskalafagadyleqdlamtkdgrlv 61

QY 64 VTHDHFGLDGLTDAVAKFPFHRHRKDGYYVIDFTLKEIQSLEMTENFETMGKWSKSSVYG 123  
 Db 62 vihdhflldgltdavakffphrhrkdgryyvidftlkeiqslemtentfmggkwsksvvg 121

QY 124 WPTVTRMRRAEPAADGVGAASRDLEKKGAISSNTAATNAACAWLEAQEEVEEGFPVTP 183  
 Db 122 wptvtrmreraepaadvgaasrdlekgaitssntaataaacawleaqeeveegfpvtp 181

QY 184 QVPLRPMTYKAADVLSHFLKEKGLLEGLIHSORRQDILDILWIYHTGYPDPQNTYTPGPG 243  
 Db 182 qvplrpmtykaadvlsfhflkekgleglilhsgrrqddildilwiyhtgypdqnytpgpg 241

QY 244 VRYPTFTFGCYKLVPEPKVEEANKGENTSLHHPVSLHGMDDPEREVLWRFDSSLAFH 303  
 Db 242 vryptftfgwcyklvpvpdkveeankgentsllhpsvlhgmddperevlewrfdslafh 301

QY 304 HVARELHPEYFKNCTSEPVDPRLPEPMKHPGSPKTACTIONCYCKKCFHCQVCFITKALGI 363  
 Db 302 hvarelhpeyfknc-----tsghhhhhh 315

QY 364 SYGRKKRRRRRPPQGSQTHQVSLSKQPTSQSRGDPGPKETSGHHHHH 413  
 Db 316 -----tsghhhhhh 324

RESULT 8  
 AAG63234  
 ID AAG63234 standard; Protein; 324 AA.  
 XX AC AAG63234;  
 DT 01-OCT-2001 (first entry)  
 DE Amino acid sequence of a His tagged LipD-Nef of HIV.  
 KW HIV; nef gene; tat gene; vaccine; gp120 gene; HIV viral load.  
 XX Synthetic.  
 OS Human immunodeficiency virus.  
 XX Key Location/Qualifiers  
 FH Peptide 1..109  
 FT /note= "ProtD fusion partner"  
 XX WO200154719-A2.  
 XX 02-AUG-2001.  
 XX 29-JAN-2001; 2001WO-EP00944.  
 XX 31-JAN-2000; 2000GB-0002200.  
 PR 14-APR-2000; 2000GB-0009336.  
 PR 06-JUN-2000; 2000GB-0013806.  
 PR 28-JUN-2000; 2000WO-EP05998.  
 XX (SMIK ) SMITHKLINE BEECHAM BIOLOGICALS.  
 XX Voss G;  
 XX WPI; 2001-476172/51.  
 DR N-PSDB; AAX42878.  
 XX New use of a human immunodeficiency virus (HIV) Tat, or Nef, or Nef

PT linked to Tat (Nef-Tat) protein or polynucleotide and an HIV gp120  
 XX protein or polynucleotide for the manufacture of a vaccine  
 PS Disclosure; Fig 1; 90pp; English.

CC The present sequence represents a His-tagged Nef protein of HIV, with  
 CC a lipidation signal sequence (LipoD) which is removed after processing.  
 CC The protein is expressed in *Escherichia coli*, and is used to  
 CC produce the vaccine of the invention. The specification describes  
 CC the use of HIV Tat, HIV Nef, or Nef-Tat; and HIV gp120 in the  
 CC manufacture of a vaccine. The vaccine is used for the prophylactic or  
 CC therapeutic immunization of humans against HIV. Tat, Nef or Nef-Tat act  
 CC in synergy with gp120 in the treatment and prevention of HIV. The  
 CC vaccine reduces the HIV viral load in HIV infected humans and results  
 CC in a maintenance of CD4+ levels over those levels found in the absence  
 CC of vaccination with HIV Tat, Nef or Nef-Tat and HIV gp120.

SX Sequence 324 AA;

Query Match 75.0%; Score 1697.5; DB 22; Length 324;  
 Best Local Similarity 78.8%; Pred. No. 7.5e-152;  
 Matches 323; Conservative 0; Mismatches 0; Indels 87; Gaps 1;

QY 4 SSSSSNANTOMKSDKIIIIAHGASGYLPHTLESKALAFQAQADYLEQDLAMTKOGRV 63  
 DB 2 SSSSSNANTOMKSDKIIIIAHGASGYLPHTLESKALAFQAQADYLEQDLAMTKOGRV 61  
 QY 64 VIHDFLDGLTDVAKKPHRHRDGRYYVDFTLKEIQSLEMTENFTMGKWSKSSVVG 123  
 DB 62 VIHDFLDGLTDVAKKPHRHRDGRYYVDFTLKEIQSLEMTENFTMGKWSKSSVVG 121  
 QY 124 WPTVRMRRAEPAADGVGAASRDLEKHGAITSNTAATNAACAWLEAQEEVEGFPVTP 183  
 DB 122 WPTVRMRRAEPAADGVGAASRDLEKHGAITSNTAATNAACAWLEAQEEVEGFPVTP 181  
 QY 184 QVLRPMTYKAAVDLSHFLKEKGLGLIHSQRRQDILDLYHTQGYFPDQWNTPGPG 243  
 DB 182 QVLRPMTYKAAVDLSHFLKEKGLGLIHSQRRQDILDLYHTQGYFPDQWNTPGPG 241  
 QY 244 VRYPLTGWYKLVVPEDKVEEANKGENTSLHPVSLHGMDPDEREVRDLSRLAFH 303  
 DB 242 VRYPLTGWYKLVVPEDKVEEANKGENTSLHPVSLHGMDPDEREVRDLSRLAFH 301  
 QY 304 HVARELHPEYFKNCTSEPVDRPLEPKHGPSQPKTACTNCKKCCFHCQVCFITKALGI 363  
 DB 302 HVARELHPEYFKNCTSEPVDRPLEPKHGPSQPKTACTNCKKCCFHCQVCFITKALGI 315  
 QY 364 SYGRKKRRRRRPPQGSQTHQVSLSKQPTSGSRGDPGPKETSGHHHHH 413  
 DB 316 SYGRKKRRRRRPPQGSQTHQVSLSKQPTSGSRGDPGPKETSGHHHHH 324

RESULT 9

AAY02351  
 ID AAY02351 standard; Protein; 302 AA.

XX AC AAY02351;

XX DT 09-JUL-1999 (first entry)

XX DE A representative HIV-1 Nef-Tat-His protein.

XX KW HIV nef gene; fusion protein; HIV nef gene; Nef protein; Tat protein;  
 XX KW vaccine; HIV infection; protein D.

XX OS Synthetic.

XX OS Human immunodeficiency virus type 1.

XX PN WO9916884-A1.

XX PD 08-APR-1999.

XX XX

PF 17-SEP-1998; 98WO-EP06040.  
 XX  
 PR 26-SEP-1997; 97GB-0020585.  
 XX  
 XX (SMK ) SMITHKLINE BEECHAM BIOLOGICALS.  
 PA

XX Bruck C, Godart SAG, Marchand M;

XX WPI; 1999-302282/25.

DR N-PSDB; AAX35687.

XX HIV Tat or Nef protein linked to a fusion partner

XX Disclosure; Fig 2; 66pp; English.

CC The present sequence represents a representative HIV-1 Nef-Tat-His  
 CC protein. The protein is used in the creation of the fusion proteins of  
 CC the invention, in conjunction with a fusion partner (e.g. protein D).  
 CC The fusion protein can be used in a vaccine to prevent HIV infection.

SX Sequence 302 AA;

Query Match 74.6%; Score 1688; DB 20; Length 302;  
 Best Local Similarity 100.0%; Pred. No. 5.4e-151;  
 Matches 302; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 112 MGKWSKSSVVGWPTVRMRRAEPAADGVGAASRDLEKHGAITSNTAATNAACAWLEA 171  
 DB 1 MGKWSKSSVVGWPTVRMRRAEPAADGVGAASRDLEKHGAITSNTAATNAACAWLEA 60  
 QY 172 QEEVEYGFVTPQVLRPMTYKAAVDLSHFLKEKGLGLIHSQRRQDILDLYHTQGY 231  
 DB 61 QEEVEYGFVTPQVLRPMTYKAAVDLSHFLKEKGLGLIHSQRRQDILDLYHTQGY 120  
 QY 232 FPDQWNTPGGYRPLTFGWYKLVVPEDKVEEANKGENTSLHPVSLHGMDPDERE 291  
 DB 121 FPDQWNTPGGYRPLTFGWYKLVVPEDKVEEANKGENTSLHPVSLHGMDPDERE 180  
 QY 292 LEVRFSRLAFHVAARELHPEYFKNCTSEPVDRPLEPKHGPSQPKTACTNCKKCCFH 351  
 DB 181 LEVRFSRLAFHVAARELHPEYFKNCTSEPVDRPLEPKHGPSQPKTACTNCKKCCFH 240  
 QY 352 QVCFITKALGISYGRKKRRRRRPPQGSQTHQVSLSKQPTSGSRGDPGPKETSGHHH 411  
 DB 241 QVCFITKALGISYGRKKRRRRRPPQGSQTHQVSLSKQPTSGSRGDPGPKETSGHHH 300  
 QY 412 HH 413  
 DB 301 HH 302

RESULT 10

AAG63233  
 ID AAG63233 standard; Protein; 302 AA.

XX AC AAG63233;

XX DT 01-OCT-2001 (first entry)

XX DE Amino acid sequence of a His-tagged Nef-Tat linked protein of HIV.

XX KW HIV; nef gene; tat gene; vaccine; gp120 gene; HIV viral load.

XX OS Synthetic.

XX OS Human immunodeficiency virus.

XX PN WO200154719-A2.

XX PD 02-AUG-2001.

XX PF 29-JAN-2001; 2001WO-EP00944.

XX XX

PR 31-JAN-2000; 2000GB-0002200.  
PR 14-APR-2000; 2000GB-0009336.  
PR 06-JUN-2000; 2000GB-0013806.  
PR 28-JUN-2000; 2000WO-EP05998.  
XX (SMIK ) SMITHKLINE BEECHAM BIOLOGICALS.  
XX PA  
XX PI Voss G;  
XX  
XX WPI; 2001-476172/51.  
DR N-PSDB; AAH42877.  
XX  
XX New use of a human immunodeficiency virus (HIV) Tat, or Nef, or Nef  
PT linked to Tat (Nef-Tat) protein or polynucleotide and an HIV gp120  
PT protein or Polynucleotide for the manufacture of a vaccine -  
XX  
XX Disclosure: Fig 1; 90pp; English.  
XX  
XX The present sequence represents a His-tagged Nef-Tat linked protein of  
CC HIV. The protein is expressed in the yeast Pichia pastoris, and is used  
CC to produce the vaccine of the invention. The specification describes  
CC the use of HIV Tat, HIV Nef, or Nef-Tat; and HIV gp120 in the  
CC manufacture of a vaccine. The vaccine is used for the prophylactic or  
CC therapeutic immunization of humans against HIV. Tat, Nef or Nef-Tat act  
CC in synergy with gp120 in the treatment and prevention of HIV. The  
CC vaccine reduces the HIV viral load in HIV infected humans and results  
CC in a maintenance of CD4+ levels over those levels found in the absence  
CC of vaccination with HIV Tat, Nef or Nef-Tat and HIV gp120.  
XX  
XX Sequence 302 AA;  
SQ

Query Match 74.6%; Score 1688; DB 22; Length 302;  
Best Local Similarity 100.0%; Pred. No. 5.4e-151;  
Matches 302; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
QY 112 MGGKWSKSVGWGPTVRERMRRAEPAADGVGAASRDLEKHGAITSNTAATNAACAWLEA 171  
Db 1 mggkwsksvsgvptvrermraraepaadvgaasrdlekhgaitssntaatnaacawlea 60  
QY 172 QEEVEGFPVTPQVPLRPMYTYKAAVDLSHFLKEKGLGLIHSQRRQDILDLIWYHTQGY 231  
Db 61 qeeevgfpvtpqvpplrmpntykaavdlshflkekggleglihsqrqrqldliwytggy 120  
QY 232 FPDWQNTYTPGCVRYPLTFGWCKYKLVPEPKVEANKGENTSLHPVSLHGMDDPEREV 291  
Db 121 fpdwnqytpggyrvyplttfgwcyklyvppepkveeankgentsllhpvslhgmddperev 180  
QY 292 LEWRDLSRLAFHHVARELHPEYFKNCTSEPDPRLEPKHGPSQPKTACTNCCYKCCCFH 351  
Db 181 lewrdsrlafhhvarelhpeyfkntsepdprlepkhpgsqpkactnccyckcccfh 240  
QY 352 CQVCFITKALGISYGRKKRRRRPPQGSQTHQVSLSKOPTSQSRGDPGPKETSGHHHH 411  
Db 241 cqvcfitaalgisygrkkrrrrppqgsqthqvslskptsqsgsdgdpkpketsgghhhh 300  
QY 412 HH 413  
Db 301 hh 302

RESULT 11  
AAV02357  
ID AAV02357 standard; Protein; 302 AA.  
XX AC AAV02357;  
XX  
XX 09-JUL-1999 (first entry)  
XX  
XX A representative HIV-1 mutant Tat-His protein.  
XX  
XX HIV nef gene; fusion protein; HIV nef gene; Nef protein; Tat protein;  
KW vaccine; HIV infection; protein D.

XX Synthetic.  
OS Human immunodeficiency virus type 1.  
XX  
XX WO9916884-A1.  
PN  
XX 08-APR-1999.  
PD  
XX  
XX 17-SEP-1998; 98WO-EP06040.  
PF  
XX 26-SEP-1997; 97GB-0020585.  
PR  
XX (SMIK ) SMITHKLINE BEECHAM BIOLOGICALS.  
PA  
XX Bruck C, Godart SAG, Marchand M;  
PI  
XX WPI; 1999-302282/25.  
DR N-PSDB; AAX35693.  
XX  
XX HIV Tat or Nef protein linked to a fusion partner  
PT  
XX Disclosure: Fig 2; 66pp; English.  
PS  
XX The present sequence represents a representative HIV-1 mutant Tat-His  
CC protein. The protein is used in the creation of the fusion proteins of  
CC the invention, in conjunction with a fusion partner (e.g. protein D).  
CC The specification also describes fusion proteins comprising HIV-1 Nef  
CC protein. The fusion protein can be used in a vaccine to prevent HIV  
CC infection.  
XX  
XX Sequence 302 AA;  
SQ

Query Match 74.0%; Score 1675; DB 20; Length 302;  
Best Local Similarity 99.0%; Pred. No. 9.1e-150;  
Matches 299; Conservative 2; Mismatches 1; Indels 0; Gaps 0;  
QY 112 MGGKWSKSVGWGPTVRERMRRAEPAADGVGAASRDLEKHGAITSNTAATNAACAWLEA 171  
Db 1 mggkwsksvsgvptvrermraraepaadvgaasrdlekhgaitssntaatnaacawlea 60  
QY 172 QEEVEGFPVTPQVPLRPMYTYKAAVDLSHFLKEKGLGLIHSQRRQDILDLIWYHTQGY 231  
Db 61 qeeevgfpvtpqvpplrmpntykaavdlshflkekggleglihsqrqrqldliwytggy 120  
QY 232 FPDWQNTYTPGCVRYPLTFGWCKYKLVPEPKVEANKGENTSLHPVSLHGMDDPEREV 291  
Db 121 fpdwnqytpggyrvyplttfgwcyklyvppepkveeankgentsllhpvslhgmddperev 180  
QY 292 LEWRDLSRLAFHHVARELHPEYFKNCTSEPDPRLEPKHGPSQPKTACTNCCYKCCCFH 351  
Db 181 lewrdsrlafhhvarelhpeyfkntsepdprlepkhpgsqpkactnccyckcccfh 240  
QY 352 CQVCFITKALGISYGRKKRRRRPPQGSQTHQVSLSKOPTSQSRGDPGPKETSGHHHH 411  
Db 241 cqvcfitaalgisygrkkrrrrppqgsqthqvslskptsqsgsdgdpkpketsgghhhh 300  
QY 412 HH 413  
Db 301 hh 302

RESULT 12  
AAG63239  
ID AAG63239 standard; Protein; 302 AA.  
XX  
XX AAG63239;  
XX  
XX 01-OCT-2001 (first entry)  
DT  
XX Amino acid sequence of a His-tagged mutant His protein of HIV.  
DE  
XX HIV; nef gene; tat gene; vaccine; gp120 gene; HIV viral load.  
KW

XX Synthetic.  
OS Human immunodeficiency virus.  
XX WO200154719-A2.  
XX 02-AUG-2001.  
XX 29-JAN-2001; 2001WO-EP00944.  
XX 31-JAN-2000; 2000GB-0002200.  
XX 14-APR-2000; 2000GB-0009336.  
XX 06-JUN-2000; 2000GB-0013806.  
XX 28-JUN-2000; 2000WO-EP05998.  
XX (SMIK ) SMITHKLINE BEECHAM BIOLOGICALS.  
XX Voss G;  
XX WPI; 2001-476172/51.  
XX N-PSDB; AAH42883.  
XX New use of a human immunodeficiency virus (HIV) Tat, or Nef, or Nef  
XX linked to Tat (Nef-Tat) protein or polynucleotide and an HIV gp120  
XX protein or polynucleotide for the manufacture of a vaccine .  
XX Disclosure; Fig 1; 90pp; English.  
XX The present sequence represents a His-tagged mutant His protein of HIV.  
XX The protein is expressed in the yeast Pichia pastoris, and is used to  
XX produce the vaccine of the invention. The specification describes  
XX the use of HIV Tat, HIV Nef, or Nef-Tat; and HIV gp120 in the  
XX manufacture of a vaccine. The vaccine is used for the prophylactic or  
XX therapeutic immunization of humans against HIV. Tat, Nef or Nef-Tat act  
XX in synergy with gp120 in the treatment and prevention of HIV. The  
XX vaccine reduces the HIV viral load in HIV infected humans and results  
XX in a maintenance of CD4+ levels over those levels found in the absence  
XX of vaccination with HIV Tat, Nef or Nef-Tat and HIV gp120.  
XX Sequence 302 AA;

Query Match 74.0%; Score 1675; DB 22; Length 302;  
Best Local Similarity 99.0%; Pred. No. 9.1e-150;  
Matches 299; Conservative 2; Mismatches 1; Indels 0; Gaps 0;  
QY 112 MGGKWSKSVVGVPTVVRERMRAEPAADGVAASRDLEKHGAITSNTAATNAACAWLEA 171  
DB 1 mggkwsksvvgvptvvrermraeapaadgvaasrdlekhgaitsntaataacawlea 60  
QY 172 QEEVEGFPVTPQVPLRPMTYKAAVDLSHFLKEKGLEGLIHSQRQDILDWYHTQGY 231  
DB 61 qeeevgfpvtpqvlrpmtykaavdlshflkekggleglilhsqrqrqldilwlyhtggy 120  
QY 232 FPDWQNYTPGCVRYPLTFGCYKLVPEPKVEEANKGENTSLHPVSLHGMDDPEREV 291  
DB 121 fpdwnytpgpgvrypltfgcycylvppepkveeankgentsllhpvslhgmddperev 180  
QY 292 LEWRDLSRLAFHHVARELHPEYFKNCTSEPDVPRLEPKHKGSPQKTACTIONCYCKCCFPH 351  
DB 181 lewrfdslafhhvarelhpeyfkntsepdvprlepkhkgspqktaactioncyckccfph 240  
QY 352 CQVCFITKALGISYGRKKRRRRPPQGSQTHVSLSKOPTSOSRGDPTGPKETSGHHHH 411  
DB 241 cqvcfitaalgisygrkkrrrrppqgsqthvslskoptsqskgeptgsketsghhhh 300  
QY 412 HH 413  
DB 301 hh 302  
RESULT 13  
AAY02349

ID AAY02349 standard; Protein; 215 AA.  
XX AAY02349;  
XX 09-JUL-1999 (first entry)  
XX A representative HIV-1 Nef-His protein.  
XX HIV nef gene; fusion protein; HIV nef gene; Nef protein; Tat protein;  
XX vaccine; HIV infection; protein D.  
XX Synthetic.  
XX Human immunodeficiency virus type 1.  
XX WO9916884-A1.  
XX 08-APR-1999.  
XX 17-SEP-1998; 98WO-EP06040.  
XX 26-SEP-1997; 97GB-0020585.  
XX (SMIK ) SMITHKLINE BEECHAM BIOLOGICALS.  
XX Bruck C, Godart SAG, Marchand M;  
XX WPI; 1999-302282/25.  
XX N-PSDB; AAX35685.  
XX HIV Tat or Nef protein linked to a fusion partner  
XX Disclosure; Fig 2; 66pp; English.  
XX The present sequence represents a representative HIV-1 Nef-His protein.  
XX The protein is used in the creation of the fusion proteins of  
XX the invention, in conjunction with a fusion partner (e.g. protein D).  
XX The specification also describes fusion proteins comprising HIV-1 Tat  
XX protein. The fusion protein can be used in a vaccine to prevent HIV  
XX infection.  
XX Sequence 215 AA;

Query Match 50.3%; Score 1139.5; DB 20; Length 215;  
Best Local Similarity 71.2%; Pred. No. 2.3e-99;  
Matches 215; Conservative 0; Mismatches 0; Indels 87; Gaps 1;  
QY 112 MGGKWSKSVVGVPTVVRERMRAEPAADGVAASRDLEKHGAITSNTAATNAACAWLEA 171  
DB 1 mggkwsksvvgvptvvrermraeapaadgvaasrdlekhgaitsntaataacawlea 60  
QY 172 QEEVEGFPVTPQVPLRPMTYKAAVDLSHFLKEKGLEGLIHSQRQDILDWYHTQGY 231  
DB 61 qeeevgfpvtpqvlrpmtykaavdlshflkekggleglilhsqrqrqldilwlyhtggy 120  
QY 232 FPDWQNYTPGCVRYPLTFGCYKLVPEPKVEEANKGENTSLHPVSLHGMDDPEREV 291  
DB 121 fpdwnytpgpgvrypltfgcycylvppepkveeankgentsllhpvslhgmddperev 180  
QY 292 LEWRDLSRLAFHHVARELHPEYFKNCTSEPDVPRLEPKHKGSPQKTACTIONCYCKCCFPH 351  
DB 181 lewrfdslafhhvarelhpeyfkntsepdvprlepkhkgspqktaactioncyckccfph 206  
QY 352 CQVCFITKALGISYGRKKRRRRPPQGSQTHVSLSKOPTSOSRGDPTGPKETSGHHHH 411  
DB 207 -----tsghhhh 213  
QY 412 HH 413  
DB 214 hh 215  
RESULT 14



AAG63231

ID AAG63231 standard; Protein; 215 AA.

XX AAG63231;

XX 01-OCT-2001 (first entry)

XX Amino acid sequence of a His-tagged Nef protein of HIV.

KW HIV; nef gene; tat gene; vaccine; gp120 gene; HIV viral load.

XX Synthetic.

OS Human immunodeficiency virus.

XX WO200154719-A2.

XX 02-AUG-2001.

XX 29-JAN-2001; 2001WO-EP00944.

XX 31-JAN-2000; 2000GB-0002200.

PR 14-APR-2000; 2000GB-0009336.

PR 06-JUN-2000; 2000GB-0013806.

PR 28-JUN-2000; 2000WO-EP05998.

XX (SMIK) SMITHKLINE BEECHAM BIOLOGICALS.

XX Voss G;

XX WPI; 2001-476172/51.

DR N-PSDB; AAH42875.

XX New use of a human immunodeficiency virus (HIV) Tat, or Nef, or Nef  
PT linked to Tat (Nef-Tat) protein or polynucleotide and an HIV gp120  
PT protein or polynucleotide for the manufacture of a vaccine -

XX Disclosure; Fig 1; 90pp; English.

XX The present sequence represents a His-tagged Nef protein of HIV. The  
CC protein is expressed in the yeast *Pichia pastoris*, and is used to  
CC produce the vaccine of the invention. The specification describes  
CC the use of HIV Tat, HIV Nef, or Nef-Tat; and HIV gp120 in the  
CC manufacture of a vaccine. The vaccine is used for the prophylactic or  
CC therapeutic immunization of humans against HIV. Tat, Nef or Nef-Tat act  
CC in synergy with gp120 in the treatment and prevention of HIV. The  
CC vaccine reduces the HIV viral load in HIV infected humans and results  
CC in a maintenance of CD4+ levels over those levels found in the absence  
CC of vaccination with HIV Tat, Nef or Nef-Tat and HIV gp120.

XX Sequence 215 AA:

Query Match 50.3%; Score 1139.5; DB 22; Length 215;  
Best Local Similarity 71.2%; Pred. No. 2.3e-99;  
Matches 215; Conservative 0; Mismatches 0; Indels 87; Gaps 1;

QY 112 MGGKWSKSVVGPVTRMRRAEPAADGVGAASRDLEKHGAITSNTAATNAACAWLEA 171

Db 1 mggkwsksvvgpvrermraepaadvgaasrdlekhgaitsntaataacawlea 60

QY 172 QEEVEVGFPTVPQVPLRPMTYKAAVDLSHFLKEGGLIHSQRQDILDLMVHTOGY 231

Db 61 qeevevgfvtpqvpplrpmtykaavdlshflkekggleghsqrqdlldliwhtggy 120

QY 232 FPDWQNYTPGQVRYPLTFGCYKLVPEPKVEANKGENTSLHPVSLHGMDDPEREV 291

Db 121 fpdwnytpgpgvrypltfgcycyklvppepkveeankgentsllhpsvslhgmddperev 180

QY 292 LEWRDLSRLAFHHVARELHPEYFKNC 351

Db 181 lewrfdslrlafhhvarelhpeyfknc----- 206

QY 352 CQVCFTITRALGISYGRKKRRRRRPPQSGQTHVSLSKQPTQSQRCDPTGPKETSGHHH 411

Db 207 -----|||||

QY 412 RH 413

Db 214 hh 215

Db 214 hh 215

RESULT 15

AAY50795

ID AAY50795 standard; protein; 206 AA.

XX AAY50795;

XX 17-FEB-2000 (first entry)

XX Human NEF protein/calmodulin binding inhibitor.

XX NEF protein; calmodulin; negative factor protein; binding inhibitor;  
KW diagnosis; detection; infection; treatment; HIV.

XX Homo sapiens.

XX WO957136-A2.

XX 11-NOV-1999.

XX 06-MAY-1999; 99WO-EP03105.

XX 06-MAY-1998; 98DE-1020224.

XX (SCHO/) SCHOTT M.

XX (SCHO/) SCHORR J.

XX (ANTZ/) ANTZ C.

XX Schott M, Schorr J, Antz C;

XX WPI; 2000-038789/03.

XX Binding agents used for treatment, prevention and diagnosis of human  
PT immune deficiency virus infection -

XX Claim 3; Fig 1; 29pp; German.

XX This invention describes novel binding partners (A) (i) for negative  
CC factor protein (Nef) that competitively inhibit binding of calmodulin (i)  
CC to Nef, or (ii) for (i) that competitively inhibit binding of Nef to (i).  
CC (A) are used for prevention, diagnosis (by specific detection of intra-  
CC and/or extra-cellular Nef, including staging of infection), and/or  
CC treatment of infections with human immune deficiency virus (HIV).

XX Sequence 206 AA:

Query Match 49.5%; Score 1120; DB 21; Length 206;

Best Local Similarity 99.0%; Pred. No. 1.5e-97;

Matches 204; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 112 MGGKWSKSVVGPVTRMRRAEPAADGVGAASRDLEKHGAITSNTAATNAACAWLEA 171

Db 1 mggkwsksvvgpvrermraepaadvgaasrdlekhgaitsntaanaacawlea 60

QY 172 QEEVEVGFPTVPQVPLRPMTYKAAVDLSHFLKEGGLIHSQRQDILDLMVHTOGY 231

Db 61 qeevevgfvtpqvpplrpmtykaavdlshflkekggleghsqrqdlldliwhtggy 120

QY 232 FPDWQNYTPGQVRYPLTFGCYKLVPEPKVEANKGENTSLHPVSLHGMDDPEREV 291

Db 121 fpdwnytpgpgvrypltfgcycyklvppepkveeankgentsllhpsvslhgmddperev 180

QY 292 LEWRDLSRLAFHHVARELHPEYFKNC 317

Db 181 lewrfdslrlafhhvarelhpeyfknc 206

## RESULT 16

AAR38893  
ID AAR38893 standard; Protein; 206 AA.XX AC AAR38893;  
XX DT 10-NOV-1993 (first entry)  
XX DE Nef protein of HIV-1.  
XX KW AIDS; antibody; p25; gp110; gp41; assay; detection;  
XX OS immunity; vaccine.

Human immunodeficiency virus-1.

XX OS US5221610-A.

XX PN 22-JUN-1993.

XX PF 26-MAY-1988; 88US-0199143.

XX PR 26-MAY-1988; 88US-0199143.

XX PR 04-SEP-1991; 91US-0754300.

XX PA (INRM ) INST NAT SANTE & RECH MEDICALE.  
XX PA (INSP ) INST PASTEUR.XX PI Bahraoui EM, Chamaret S, Ferris S, Granier C, Montagnier L;  
XX PI Rietschoten JV, Rochat H, Sabatier JM;  
XX DR WPI; 1993-213434/26.XX Diagnosis of HIV infection - by detecting HIV antibodies using  
XX antigenic polypeptide derived from nef protein of HIV-1

XX Disclosure; Fig 2; 15pp; English.

XX The nef protein comprises peptides which are expressed in vivo in HIV  
XX infected patients before detectable ams. of p25, gp110 and gp41 are  
XX expressed. Thus, they can be used in assays for early detection of HIV.  
XX They can also be used to raise antibodies for use in detection.  
XX to induce cellular immunity or to raise neutralising antibodies  
XX that either inactivate the AIDS virus or reduce the viability of  
XX the virus in vivo or destroy infected cells.  
XX The peptides may be used in viral vaccines.

XX SQ Sequence 206 AA;

Query Match 49.3%; Score 1116; DB 14; Length 206;  
Best Local Similarity 99.0%; Pred. No. 3.5e-97;  
Matches 204; Conservative 0; Mismatches 2; Indels 0; Gaps 0;QY 112 MGGKWSKSSVVGWPTVRMRRAEPAADGVGAASRDLEKHGAITSNTAATNAACAWLEA 171  
DB 1 mggkwsksvvgwptvrermraepaadvgaasrdlekhgaitssntaataacawlea 60  
QY 172 QEEVEVGFPTVPOVPLRPMTYKAAVDLSHFLKEKGGLEGLIHSQRQDILDWYHTOGY 231  
DB 61 qeevevgfptvpqvpplrpmtykaavdlshflkekggleglshqrqddildwiyhtggy 120  
QY 232 FPDWQNYTPGCGVRYPLTFGWCYKLVPEPKVEANKGENTSLHLPVSLHGMDDPEREV 291  
DB 121 fpdwqnytpgpgvrypltfwgcyklvppepkveeankgentsllhpvsihgmddperev 180  
QY 292 LEWRDSDSLAFHHVARELHPEYFKNC 317  
DB 181 lewrdsrlafhhvarelhpeyfknc 206

## RESULT 17

AAB10054

ID AAB10054 standard; Protein; 206 AA.

XX AC AAB10054;  
XX DT 02-NOV-2000 (first entry)

XX DE HIV-1 nef protein.

XX KW Glycoprotein; gag; pol; GP-1; GP-2; anti-HIV; cytostatic;  
XX KW gene therapy; treatment; infectious disease; HIV; AIDS; neoplasm;  
XX KW carcinoma; melanoma; nef protein.

XX OS Human immunodeficiency virus type 1.

XX PN EP1006196-A2.

XX PD 07-JUN-2000.

XX PF 25-NOV-1999; 99EP-0250415.

XX PR 26-NOV-1998; 98DE-1056463.

XX PA (PETT-) PETTE INST HEINRICH.

XX PI Von Laer MD;

XX DR WPI; 2000-378268/33.

XX DR N-PSDB; AAA40298, AAB10053, AAB10054.

XX PT New retroviral packing cell useful as pharmaceutical carrier in gene  
XX therapy for treatment of HIV and neoplasms, comprises retroviral genes  
XX PT and glycoproteins -

XX PS Disclosure; Page 48; 69pp; German.

XX This invention describes a novel retroviral packing cell (I), comprising  
XX the retroviral genes gag, pol and glycoproteins GP-1 and GP-2 of the LCMV  
XX coding gene gp, or a part of these. The products of the invention have  
XX anti-HIV and cytostatic activity and can be used for gene therapy. (I) is  
XX useful for in vitro infection of cells, especially hematopoietic stem  
XX cells, for expression of transgenes in cells and as a pharmaceutical  
XX carrier for gene therapy. (I) is therefore useful in the treatment of  
XX infectious diseases (HIV or AIDS), neoplasms (carcinomas), melanomas and  
XX other diseases. This sequence represents the Human immunodeficiency virus  
XX (HIV-1) nef protein described in the method of the invention.

XX SQ Sequence 206 AA;

Query Match 49.2%; Score 1115; DB 21; Length 206;  
Best Local Similarity 98.1%; Pred. No. 4.4e-97;  
Matches 202; Conservative 1; Mismatches 3; Indels 0; Gaps 0;QY 112 MGGKWSKSSVVGWPTVRMRRAEPAADGVGAASRDLEKHGAITSNTAATNAACAWLEA 171  
DB 1 mggkwsksvvgwptvrermraepaadvgaasrdlekhgaitssntaataacawlea 60  
QY 172 QEEVEVGFPTVPOVPLRPMTYKAAVDLSHFLKEKGGLEGLIHSQRQDILDWYHTOGY 231  
DB 61 qeevevgfptvpqvpplrpmtykaavdlshflkekggleglshqrqddildwiyhtggy 120  
QY 232 FPDWQNYTPGCGVRYPLTFGWCYKLVPEPKVEANKGENTSLHLPVSLHGMDDPEREV 291  
DB 121 fpdwqnytpgpgvrypltfwgcyklvppepkveeankgentsllhpvsihgmddperev 180  
QY 292 LEWRDSDSLAFHHVARELHPEYFKNC 317  
DB 181 lewrdsrlafhhvarelhpeyfknc 206\*RESULT 18  
AAW89326

```

ID AAW89326 standard; Protein; 206 AA.
XX
AC AAW89326;
XX
DT 01-JUN-1999 (first entry)
XX
DE HIV-1 nef protein sequence.
XX
KW Antigenic composition; primate; lentivirus; nef gene; vaccine;
infection; AIDS; HIV-1; nef protein.
XX
OS Human immunodeficiency virus type 1.
XX
PN US851813-A.
XX
PD 22-DEC-1998.
XX
PF 27-JAN-1994; 94US-0188583.
XX
PR 27-JAN-1994; 94US-0188583.
PR 12-JUL-1990; 90US-0551945.
PR 09-JUL-1991; 91US-0727494.
XX
PA (HARD ) HARVARD COLLEGE.
XX
PI Destosiers RC;
XX
XX WPI; 1999-080408/07.
DR N-PSDB; AAW81866.
XX
XX Lentivirus antigenic compositions - containing lentivirus with nef
PT gene deletion
XX
PS Disclosure; Fig 2A-R; 93pp; English.
XX
XX The invention relates to an antigenic composition comprising an isolated
CC primate lentivirus whose genome contains an engineered non-reversible
CC null mutation in the nef gene, or an infectious DNA clone in a carrier.
CC The antigenic composition is used in vaccines against infection by the
CC lentivirus, e.g. AIDS.
XX
SQ Sequence 206 AA;

Query Match 49.2%; Score 1114; DB 20; Length 206;
Best Local Similarity 97.6%; Pred. No. 5.4e-97;
Matches 201; Conservative 3; Mismatches 2; Indels 0; Gaps 0;

QY 112 MGGKWSKSSVVGWPTVRRMRRAEPAADGVGAASRDLEKHGALTSSNTAATNAACAWLEA 171
Db 1 mggkwsksvvgwptvrrmrreapadvggaasrdlekhgaltssntaannaacawlea 60

QY 172 QEEEEVGFVPTQVPLRPMTYKAAVDLSHFLKEKGLEGILHSQRRQDILDLYIHTQGY 231
Db 61 qeeekvgfvtqvpvlrmpmtyaavdlsflkekglegilhsqrrqdildlyihtqgy 120

QY 232 FPDQNYTFCGPGVRYPLTFCWCYKLVPEPDKVEEANKGENTSLHPVSLHGMDPEREV 291
Db 121 fpdwqnytpgpggirypltfwgcyklypvepdkveeankgentsllhpvslhgmdperv 180

QY 292 LEWRFDRLAFHHVARELHPEYFKNC 317
Db 181 lewrfdslafhhvarelhpeyfknc 206

RESULT 19
AAP61515
ID AAP61515 standard; Protein; 206 AA.
XX
AC AAP61515;
XX
DT 08-JUN-1991 (first entry)
XX

```

```

DE Sequence of E' protein.
XX
XX HIV; LAV; AIDS; diagnosis; vaccine.
XX
OS HTLV-IIIIB/H9 cells (ATCC CRL 8543).
XX
PN EPI87041-A.
XX
PD 09-JUL-1986.
XX
PF 23-DEC-1985; 85EP-0309454.
PR 24-DEC-1984; 84US-0685272.
PR 04-DEC-1985; 85US-0805069.
XX
PA (GETH ) GENENTECH INC.
XX
PI Capon DJ, Lasky LA;
XX
DR WPI; 1986-177602/28.
DR N-PSDB; AAN60288.
XX
XX Acquired immune deficiency syndrome polypeptide(s) ; obt'd. by
PT molecular cloning etc. and used for diagnosis and in vaccines
PT against virus disease
XX
PS Example; fig 2; 125pp; English.
XX
XX A comparison of AAN60287 with the cDNA of the HTLV-III genome
CC revealed one particular clone, designated p7.11 which contained a
CC DNA sequence encoding this peptide (AAP60308) sequence. This approx.
CC 2.2 kilobase covers the precursor gag region and encodes, 5' to 3',
CC p-12, p-15, p-24 a second p-15 protein, and approx. 300 extra base
CC pairs 3' to the gag region (see AAN60288).
XX
SQ Sequence 206 AA;

Query Match 49.1%; Score 1111; DB 7; Length 206;
Best Local Similarity 97.1%; Pred. No. 1e-96;
Matches 200; Conservative 4; Mismatches 2; Indels 0; Gaps 0;

QY 112 MGGKWSKSSVVGWPTVRRMRRAEPAADGVGAASRDLEKHGALTSSNTAATNAACAWLEA 171
Db 1 mggkwsksvvgwptvrrmrreapadvggaasrdlekhgaltssntaannaacawlea 60

QY 172 QEEEEVGFVPTQVPLRPMTYKAAVDLSHFLKEKGLEGILHSQRRQDILDLYIHTQGY 231
Db 61 qeeekvgfvtqvpvlrmpmtyaavdlsflkekglegilhsqrrqdildlyihtqgy 120

QY 232 FPDQNYTFCGPGVRYPLTFCWCYKLVPEPDKVEEANKGENTSLHPVSLHGMDPEREV 291
Db 121 fpdwqnytpgpggirypltfwgcyklypvepdkveeankgentsllhpvslhgmdperv 180

QY 292 LEWRFDRLAFHHVARELHPEYFKNC 317
Db 181 lewrfdslafhhvarelhpeyfknc 206

RESULT 20
AAP60423
ID AAP60423 standard; Protein; 216 AA.
XX
AC AAP60423;
XX
DT 20-AUG-1991 (first entry)
XX
XX Sequence of LAV virus ORF F protein.
XX
XX AIDS vaccine; diagnosis; immunoassay; HIV; HTLV-III.
XX
XX Lymphadenopathy virus.
XX

```

PN WO8602383-A.  
XX 24-APR-1986.  
XX 18-OCT-1985; 85WO-EP00548.  
XX 21-JAN-1985; 85GB-0001473.  
PR 18-OCT-1984; 84FR-0016013.  
PR 16-NOV-1984; 84GB-0029099.  
XX (CNRS ) CNRS CENT NAT RECH SCI.  
PA (INSP ) INST PASTEUR.  
XX Montagnier L, Krust B, Chamaret S, Clavel F, Chermann J-C;  
PI Barre-Sinoussi F, Allizon M, Sonigo P, Stewart C, Danos O;  
PI Wain-Hobson S;  
XX WPI: 1986-119166/18.  
DR N-PSDB; AAN60365.  
XX Purified glyco:protein and peptide(s) - are recognised by sera contg.  
FT antibodies against lymphadenopathy virus and useful in detecting  
PT AIDS antibodies or in vaccines  
XX Disclosure; Fig 4; 75pp; English.  
XX The inventors claim a polypeptide which is recognised by sera of  
CC human origin contg. antibodies against the virus of  
CC lymphadenopathies (LAV) or acquired immune deficiency syndrome  
CC (AIDS). Also claimed are various peptides corresp. to the AA  
CC sequences deducible from proteins encoded by LAV DNA, defined by  
CC specific residues (e.g. 12-32, 37-46, 49-79, 88-153) in accordance  
CC with a formula given in the specification.  
XX Sequence 216 AA:  
SQ  
Query Match 48.9%; Score 1107; DB 7; Length 216;  
Best Local Similarity 98.1%; Pred. No. 2.7e-96;  
Matches 202; Conservative 1; Mismatches 3; Indels 0; Gaps 0;  
QY 112 MGGKWSKSVVGVWPTVRMRRAEPAADGVGAASRDLEKHGAISSNTAATNAACAWLEA 471  
Db 11 mggkwsksvvgwptvrmttraepaahgvgaaspdlkghgaitssntaatnaacawlea 70  
QY 172 QEEEVGFPTVPQVPLRPMYKAAVDLSHFLKEKGLGGLHSQRQDILDWIYHTGQY 231  
Db 71 qeeevgfptvpqplrpmttykaavdlshflkekggleglshqrqrqldilwihtggy 130  
QY 232 RPDWNYTPGVRVPLTFGCYKLVPEPKVEANKGENTSLHPSVSLHGMDDPEREV 291  
Db 131 fpdwnytpgvrpyrltfgcylkvppepkveankgentsllhpsvslhgmddperrev 190  
QY 292 LEWRFSRLAFHHVARELHPYFKNC 317  
Db 191 lewrfsrlafhhvarelhpyfknc 216.  
RESULT 21  
AAW90179  
ID AAW90179 standard; Protein; 206 AA.  
XX AC AAW90179;  
XX 14-JUN-1999 (first entry)  
XX HTLV-III E' protein.  
XX AIDS; acquired immunodeficiency syndrome; viral infection; env protein;  
KW fusion polypeptide; E' protein; gag protein; pol protein; p' protein;  
KW gD signal peptide; detection; vaccination; etiological agent;  
KW infection inhibitor; AIDS associated retrovirus; p24; gp41.  
XX

OS Human lymphotropic virus type III.  
XX Key Location/Qualifiers  
FH Misc-difference 11 /label= Val, Ile  
FT Misc-difference 51 /label= Thr, Asn  
FT Misc-difference 54 /label= Asp, Ala  
FT Misc-difference 65 /label= Glu, Lys  
FT Misc-difference 151 /label= Glu, Asp  
FT Misc-difference 153 /label= Leu, Val  
XX US5853978-A.  
XX 29-DEC-1998.  
XX 29-JUL-1994; 94US-0282857.  
XX 08-MAY-1986; 86US-0861016.  
PR 24-DEC-1984; 84US-0685272.  
PR 04-DEC-1985; 85US-0805069.  
PR 02-AUG-1988; 88US-0227568.  
PR 19-NOV-1992; 92US-0979391.  
PR 29-SEP-1993; 93US-0129009.  
PR 29-JUL-1994; 94US-0282857.  
XX (GETH ) GENENTECH INC.  
XX Berman PW, Capon DJ, Lasky LA;  
PI WPI: 1999-094894/08.  
DR N-PSDB; AAV74271.  
XX New AIDS-associated fusion polypeptide(s) - used for the detection  
PT of AIDS or to inhibit infection by AIDS associated retrovirus or  
PT dissemination of such retrovirus in infected individuals  
XX Example 1; Fig 2A-J; 47pp; English.  
XX This invention describes a novel fusion polypeptide which comprises  
CC (a) a first polypeptide sequence of an AIDS associated E', env, or gag  
CC polypeptide, that specifically binds complementary antibody and (b) a  
CC second polypeptide sequence which is not an AIDS associated virus  
CC polypeptide. Also described are (1) a fusion polypeptide having a  
CC herpes simplex virus (HSV) gD signal peptide sequence fused in reading  
CC frame with a polypeptide sequence other than HSV gD, (2) a nucleic acid  
CC encoding a fusion polypeptide as in (1); (3) an expression vector  
CC comprising a nucleic acid as in (2); (4) a host comprising a vector as  
CC in (3). The AIDS-associated fusion polypeptides can be used in the  
CC detection of and vaccination against viral etiological agents of AIDS.  
CC They can also be administered as a pharmaceutical agent to inhibit  
CC infection by AIDS associated retrovirus or dissemination of such  
CC retrovirus in infected individuals.  
XX Sequence 206 AA;  
SQ  
Query Match 48.2%; Score 1092; DB 20; Length 206;  
Best Local Similarity 96.1%; Pred. No. 6.5e-95;  
Matches 198; Conservative 1; Mismatches 7; Indels 0; Gaps 0;  
QY 112 MGGKWSKSVVGVWPTVRMRRAEPAADGVGAASRDLEKHGAISSNTAATNAACAWLEA 171  
Db 1 mggkwsksvvgwptvrmttraepaahgvgaaspdlkghgaitssntaatnaacawlea 60  
QY 172 QEEEVGFPTVPQVPLRPMYKAAVDLSHFLKEKGLGGLHSQRQDILDWIYHTGQY 231  
Db 61 qeeevgfptvpqplrpmttykaavdlshflkekggleglshqrqrqldilwihtggy 120

RESULT 22

AA  
DT 10-SEP-2001 (first entry)

```
FT Modified-site 2 "Mvristylation site"
FT /note=
```

44.78: Score 1011: DB 22: Length 216:

CC is increased by a factor of 50-100 compared with the use of a vector  
CC containing the SV40 early promoter.

XX SQ Sequence 210 AA;  
Query Match 44.0%; Score 997; DB 19; Length 210;  
Best Local Similarity 85.2%; Pred. No. 6.4e-86;  
Matches 179; Conservative 16; Mismatches 11; Indels 4; Gaps 1;  
QY 112 MGGKWSKSSVVGWPTVRMR----RAEPAADGVGAASRDLEKHGAISSNTAATNAACA 167  
DB 1 mggkwsrsmggwsairermraepaadvavsdlekhgaitssntaataadca 60  
QY 168 WLEAQEEVEGFPVTPQVPLRPMTYKAAVDLSHFLKKGGLGLIHSQRQDILDLYIH 227  
DB 61 wleaqeeevgfpvpqvpplrpmtykaaldishfikekgglegliwsgrrgeildlyih 120  
QY 228 TQGYFDPQNYTPGVRYPPLTFGWCYKLVPEDPKVEEANKGENTSLHPVSLHGMDP 287  
DB 121 tqgyfdpqnytpgpylrypltfgwcklvpevekeanegenenllhpmshgmeda 180  
QY 288 EREYLEWRFDSRLAFHHVARELHPEYFKNC 317  
DB 181 ekeylvwrfdsrlafhhvarelhpeyykdc 210  
AA77299; 22-MAY-2000 (first entry)

RESULT 24  
AA77299  
ID AAY77299 standard; Protein: 210 AA.  
XX AC AAY77299;  
XX DT 22-MAY-2000 (first entry)  
XX DE HIV-1 (ATCC CRL 8597) ORF #3-encoded protein.  
XX KW HIV-1; immunoassay; antigen; pol fragment; p31; recombinant; antibody;  
XX KW detection.  
XX OS Human immunodeficiency virus type 1 'ATCC CRL 8597'.  
XX PN US6013432-A.  
XX PD 11-JAN-2000.  
XX PF 17-MAY-1995; 95US-0443434.  
XX PR 08-JUL-1993; 93US-0089407.  
XX PR 24-DEC-1987; 87US-0138894.  
XX PR 17-AUG-1992; 92US-0931154.  
XX PR 31-OCT-1984; 84US-0667501.  
XX PR 30-JAN-1985; 85US-0696534.  
XX PR 06-SEP-1985; 85US-0773447.  
XX PA (CHIR ) CHIRON CORP.  
XX PI Luciw PA, Dina D;  
XX DR WPI; 2000-170256/15.  
XX DR N-PSDB; AAZ90201.

XX PT Immunoassay for antibodies against human immune deficiency virus, for  
XX PT diagnosing infection, uses an immunogenic fragment of the pol protein  
XX PT as antigen.

XX PS Example 1; Fig 40-P; 99pp; English.

XX CC The invention relates to the improvement of HIV-1 immunoassays by the  
XX CC use of an HIV-1 antigen comprising an immunogenic fragment of  
XX CC recombinant or synthetic HIV-1 pol, which is encoded by an approximately  
XX CC 9.7 kb sequence between a BstXI restriction site at position 3006 and an  
XX CC NdeI site at position 5131 of the genome (the proviral DNA sequence is

CC given in AAZ90201). The immunogenic pol fragment is not immunologically  
CC cross-reactive with human T cell lymphotropic viruses I or II. The  
CC invention also encompasses the use of p31 as an antigen. The recombinant  
CC antigens may be produced in Escherichia coli, Saccharomyces cerevisiae or  
CC in mammalian cells. Immunoassays using the recombinant HIV proteins may  
CC be used to diagnose and stage HIV-1 infections. Sequences  
CC AAY77294-Y77299 represent proteins encoded by the genome of HIV-1 (ATCC  
CC CRL 8597).

XX SQ Sequence 210 AA;

Query Match 44.0%; Score 997; DB 21; Length 210;  
Best Local Similarity 85.2%; Pred. No. 6.4e-86;  
Matches 179; Conservative 16; Mismatches 11; Indels 4; Gaps 1;  
QY 112 MGGKWSKSSVVGWPTVRMR----RAEPAADGVGAASRDLEKHGAISSNTAATNAACA 167  
DB 1 mggkwsrsmggwsairermraepaadvavsdlekhgaitssntaataadca 60  
QY 168 WLEAQEEVEGFPVTPQVPLRPMTYKAAVDLSHFLKKGGLGLIHSQRQDILDLYIH 227  
DB 61 wleaqeeevgfpvpqvpplrpmtykaaldishfikekgglegliwsgrrgeildlyih 120  
QY 228 TQGYFDPQNYTPGVRYPPLTFGWCYKLVPEDPKVEEANKGENTSLHPVSLHGMDP 287  
DB 121 tqgyfdpqnytpgpylrypltfgwcklvpevekeanegenenllhpmshgmeda 180  
QY 288 EREYLEWRFDSRLAFHHVARELHPEYFKNC 317  
DB 181 ekeylvwrfdsrlafhhvarelhpeyykdc 210

RESULT 25  
AAE04962  
ID AAE04962 standard; Protein: 217 AA.  
XX AC AAE04962;  
XX DT 10-SEP-2001 (first entry)  
XX DE HIV-1 jrf1 Nef (G2A, L1AA) mutant protein.  
XX KW DNA vaccine: codon optimisation; Nef protein; HIV infection; anti-HIV;  
XX KW cell mediated immune response; cytotoxic T lymphocyte; CTL;  
XX KW human immunodeficiency virus-1; HIV-1; jrf1 Nef; mutant; mutein.  
XX OS Human immunodeficiency virus type 1.  
XX OS Synthetic.  
XX FH Key Location/Qualifiers  
XX FT Misc-difference 2 /note= "Wild type myristylation site at Gly2 substituted  
XX FT with Ala"  
XX FT Misc-difference 174..175 /note= "Wild type dileucine motif (Leu-Leu) substituted  
XX FT with Ala-Ala"

XX PN WO200143693-A2.

XX PD 21-JUN-2001.

XX PF 15-DEC-2000; 2000WO-US34162.

XX PR 17-DEC-1999; 99US-0172442.

XX PA (MERI ) MERCK & CO INC.

XX PI Shiver JW, Liang X, Fu T;

XX DR WPI; 2001-417878/44.

XX DR N-PSDB; AAD09605.

PT Novel DNA vaccine useful for inducing cell mediated immune response  
PT against virulent strains of HIV, comprises DNA expression vector, and  
PT DNA containing codon optimized open reading frame encoding Nef protein

PS Claim 29; Page 18; 84pp; English.

XX The present invention relates to a DNA vaccine comprising a DNA  
CC expression vector, and a DNA molecule containing a codon optimised open  
CC reading frame encoding a Nef protein or its immunogenic derivative. The  
CC Nef protein or its derivative is expressed and generates an immune  
CC response which provides a substantial level of protection against HIV  
CC infection, upon administration of to a host. The DNA vaccine is useful  
CC for inducing a cell mediated immune (cytotoxic T lymphocyte, CTL)  
CC response against infection or disease caused by virulent strains of HIV,  
CC by administering the DNA vaccine into the tissue of the vertebrate host,  
CC preferably human. It is also useful for lowering transmission rate to  
CC previously uninfected individuals and/or for reducing levels of viral  
CC loads within an infected individual, so as to prolong the asymptomatic  
CC phase of HIV infection. The present sequence is human immunodeficiency  
CC virus-1 (HIV-1) jrf1 Nef (G2A, L2AA) mutant protein. The different nef  
CC constructs of the invention are used in the exemplification of the  
CC invention.

XX Sequence 217 AA;

Query Match 44.0%; Score 996; DB 22; Length 217;  
Best Local Similarity 82.0%; Pred. No. 8.4e-86;  
Matches 178; Conservative 16; Mismatches 13; Indels 10; Gaps 1;

QY 112 MGKWSKSVWGVPVTRMRRAEPAAD-----GVGAASRDLEKKGAISSNTAA 161  
DB 1 magkwsksrvgpwtvtrmrtraepaadvrrtpeaavgvavrsrdlekhgaitssntaa 60  
QY 162 TNAACAWLEAQEEVEGFPVTPQVPLRPMYKAAVDLSHFLKEKGLEGLIHSQRQDIL 221  
DB 61 tnacawleaqeadeevgfpvtpqvlrmpmtykgavdlshflkekgleglihsqrqdil 120  
QY 222 DLWVHTQGYFDWQNYTPGPGVRYPLTFGWCYKLVPEPDKVEEANKGENTSLHHPVSL 281  
DB 121 dlwvhtqgyfdwqnytpgpgvrypltfvgwcfklypvepveeanegeencaahpmsq 180  
QY 282 HGMDOPEVLEWRFDSRLAFHHVARELHPEYFKNC 318  
DB 181 hgiedepevlewrfdskiafhvarelhpeyykdc 217

RESULT 26

AA012262  
ID AAR12262 standard; Protein; 210 AA.  
XX  
AC AAR12262;  
XX

DT 20-AUG-1991 (first entry)

DE HIV-1 strain OYI open reading frame (ORF) F protein.

XX HIV-1; AIDS; retroviruses.

XX Homo sapiens.

OS US019510-A.

XX 28-MAY-1991.

PF 28-OCT-1987; 87US-0113655.

XX 28-OCT-1987; 87US-0113655.

PA (INSP ) INST PASTEUR.

PI Wain-Hobson S, Huet T, Delaporte E, Brun-Vezinet F;

XX  
XX  
XX  
PT  
PT  
XX  
PS  
XX  
CC  
CC  
CC  
CC  
CC  
CC  
CC  
SQ

WPI; 1991-177518/24.

Purified human retrovirus - is mutant of HIV-1 having  
characteristics of HIV-1 OYI, used in diagnosis of HIV infection  
Disclosure; fig 8; 23pp; English.

This sequence constitutes the ORF F protein constituent of a new  
strain of HIV-1 retrovirus, OYI. This mutant retroviral strain is  
useful in an assay for diagnosing HIV infection. See also AAO11943  
(OYI nucleotide sequence), AAR12255-61 (other HIV OYI constituent  
proteins).

Sequence 210 AA;

Query Match 43.4%; Score 983; DB 12; Length 210;  
Best Local Similarity 82.5%; Pred. No. 1.3e-84;  
Matches 174; Conservative 22; Mismatches 9; Indels 6; Gaps 2;

QY 112 MGKWSKSVWGVPVTRMRRA-----EPAADGVGAASRDLEKKGAISSNTAATNAAC 166  
DB 1 mgkwskskmgvptirermkraelqppepaagvgaasrdlekhgaitssntaataadnc 60  
QY 167 AWLEAQEEVEGFPVTPQVPLRPMYKAAVDLSHFLKEKGLEGLIHSQRQDILDIWIY 226  
DB 61 awleaqeadeevgfpvtpqvlrmpmtykgaldishflkekglegliysqrqgdild-wvy 119  
QY 227 HQQGYFPDQNYTPGPGVRYPLTFGWCYKLVPEPDKVEEANKGENTSLHHPVSLHGMDD 286  
DB 120 hqgyfpdqnytpgpgvrypltfvgwcfklypmdpqveeanegeennsilhplshgmdd 179  
QY 287 PEREVLWRFDSRLAFHHVARELHPEYFKNC 317  
DB 180 pekevlwrfdsrlafhrmarevhpeyykdc 210

RESULT 27

AA04961  
ID AAE04961 standard; Protein; 237 AA.  
XX  
AC AAE04961;

DT 10-SEP-2001 (first entry)

DE Human tPA leader peptide-HIV-1 jrf1 Nef (opt tpanef) mutant protein.  
XX DNA vaccine; codon optimisation; Nef protein; HIV infection; anti-HIV;  
KW cell mediated immune response; cytotoxic T lymphocyte; CTL;  
KW human immunodeficiency virus-1; HIV-1; human; jrf1 Nef; opt tpanef;  
KW tissue plasminogen activator; tPA; chimeric protein; mutant; mutein.

OS Chimeric - Human immunodeficiency virus type 1.

XX Key Location/Qualifiers

FT Peptide 1..26

FT /label= Leader\_peptide

FT /note= "Human tissue plasminogen activator (tPA)

FT leader sequence"

FT Protein 27..237

FT /note= "Mature HIV-1 Nef (6-216aa) protein"

FT Domain 195..196

FT /label= Dileucine\_motif

XX WO200143693-A2.

XX 21-JUN-2001.

XX 15-DEC-2000; 2000WO-US34162.

XX 17-DEC-1999; 99US-0172442.





QY 287 PEREVLWRFDSRLAFHHVARELHPEYFKNC 317  
 ID AAB69363 standard; Protein; 206 AA.  
 Db 207 pekevlwrfasklafhhvarelhpeyykdc 237

RESULT 29  
 AAB69363  
 ID AAB69363 standard; Protein; 206 AA.  
 AC AAB69363;  
 XX  
 DT 20-APR-2001 (first entry)  
 XX  
 DE HIV-1 non-subtype B clone 962M651-8 nef protein.  
 XX  
 XX HIV-1; human immunodeficiency virus; non-subtype B; gag; pol; env;  
 KW vpr; vif; vpr; tat; rev; nef; vaccine.  
 XX  
 OS Human immunodeficiency virus type 1.  
 XX  
 PN WO200026416-A1.  
 XX  
 PD 11-MAY-2000.  
 XX  
 PF 25-OCT-1999; 99WO-US24837.  
 XX  
 PR 02-NOV-1998; 98US-0184418.  
 XX  
 PA (UABR-) UAB RES FOUND.  
 PI Hahn BH, Shaw GM, Gao F;  
 XX  
 XX WPI; 2000-365651/31.  
 DR  
 XX Novel genomic nucleic acids of non-subtype B human immunodeficiency  
 PT virus type 1 useful for detecting and treating AIDS comprises a  
 PT specific nucleotide sequence  
 XX  
 PS Claim 41; Fig 22; 131pp; English.  
 XX

The present invention provides the protein and coding sequences for a  
 number of human immunodeficiency virus (HIV) type 1 non-subtype B  
 isolates. The sequences shown include the near full-length coding  
 sequences from each isolate, and the env, pol, vif, vpr, vpu, gag, tat,  
 rev and nef proteins. These can be used to detect the presence of HIV-1  
 in a sample and to produce antibodies against non-subtype B HIV-1 virus.  
 These antibodies can be used in vaccines to prevent and treat HIV  
 infection.

XX Sequence 206 AA;  
 SQ

Query Match 42.0%; Score 952; DB 21; Length 206;  
 Best Local Similarity 81.1%; Pred. No. 1.1e-81;  
 Matches 167; Conservative 24; Mismatches 15; Indels 0; Gaps 0;

QY 112 MGKWKSSVGVGPTVRMRRAEPADGVGAASRDLEKHGALTSSNTAATNACAWLEA 171  
 Db 1 mggkwsksivgwpavrrirrtetpaagvgavsdldkygaltssntttnaacawlea 60

QY 172 QEEEEVGFPTPOVPLRPMYTKAAVDLSHFLKEKGLGGLIHSQRQDILDLYIHTQGY 231  
 Db 61 qeeeevgfprpdpvprpmtkyaaavdlshflkekglegliyskkrqelidlwlyhtqgf 120

QY 232 FPDQNYTTPGPGVRYPLTFGCWYKLVPEPDKVEANKGENTSLHPVSLHGMDPPERV 291  
 Db 121 fpdwqnytpgpgvrypltfgcwcklvpepdkveeanegenncllhpsmqggmddhrev 180

QY 292 LEWRFDSRLAFHHVARELHPEYFKNC 317  
 Db 181 lkwkfdshlahkhmarelhpeyykdc 206

RESULT 30  
 AAB86169  
 ID AAB86169 standard; Protein; 3025 AA.  
 XX  
 AC AAB86169;  
 XX  
 DT 21-AUG-2001 (first entry)  
 XX  
 DE HIV-1 subtype C protein fragment #1.  
 XX  
 XX Infection; diagnosis; human; humoral immune response; antiviral;  
 KW cellular immune response; vaccine; treatment; gene therapy.  
 XX  
 OS Human immunodeficiency virus type 1.  
 XX  
 PH Key Location/Qualifiers  
 FT Misc-difference 1..3025  
 FT /label= Xaa  
 FT /note= "Xaa represents a stop codon"  
 XX  
 PN DE10056747-A1.  
 XX  
 PD 31-MAY-2001.  
 XX  
 PF 16-NOV-2000; 2000DE-1056747.  
 XX  
 PR 16-NOV-1999; 99DE-1055089.  
 XX  
 PA (SHAO/) SHAO Y.  
 PA (GENE-) GENEART GMBH GES ANGEWANDTE BIOTECHNOLOG.  
 PI Wagner R, Wolf H, Shao Y, Graf M;  
 XX  
 XX WPI; 2001-336417/36.  
 DR N-P5DB; AAH20870.  
 XX  
 PT New nucleic acid sequences from a human immune deficiency virus  
 PT intersubtype, useful for treatment, prevention and diagnosis of  
 PT infection  
 XX  
 PS Disclosure: Fig 8A-O; 48pp; German.  
 XX

This invention describes a novel polynucleotide isolated from human  
 immunodeficiency virus type 1 subtype C/B' which can be used for the  
 induction of specific humoral and cellular immune responses. (I) and  
 polypeptides (II) encoded by them, are useful in pharmaceuticals,  
 vaccines and diagnostic agents, particularly for treatment or prevention  
 of human immune deficiency virus-1 (HIV-1) infections, also for rational  
 design of test or therapeutic reagents, or gene therapy vectors.  
 CC Polypeptides, especially antibodies, specifically directed against (II)  
 CC are similarly useful as pharmaceutical and diagnostic agents (I) are  
 CC specific for intersubtype C/B' of HIV-1 so are useful in regions  
 CC (particularly China and South-East Asia) where this subtype is prevalent.  
 CC The products of the invention have antiviral activity. This sequence  
 CC represents a protein encoded by the HIV-1 subtype C' genome described in  
 CC the method of the invention.

XX Sequence 3025 AA;  
 SQ

Query Match 41.9%; Score 948.5; DB 22; Length 3025;  
 Best Local Similarity 70.6%; Pred. No. 1.2e-79;  
 Matches 175; Conservative 22; Mismatches 42; Indels 9; Gaps 2;

QY 112 MGKWKSSVGVGPTVRMRRAEPADGVGAASRDLEKHGALTSSNTAATNACAWLEA 171  
 Db 2723 mggkwsksivgwpavrrirrtetpaagvgavsdldkygaltssntttnaacawlea 2782

QY 172 QEEEEVGFPTPOVPLRPMYTKAAVDLSHFLKEKGLGGLIHSQRQDILDLYIHTQGY 231  
 Db 2783 qeeeevgfprpdpvprpmtkyaaavdlshflkekglegliyskkrqelidlwlyhtqgy 2842

QY 232 FPDQNYTTPGPGVRYPLTFGCWYKLVPEPDKVEANKGENTSLHPVSLHGMDPPERV 291

Db 2843 fpdwnytpgpgvirfpltfgrwcklvdpvpreveeanegedncilhpvcqhgmeddhrev 2902  
QY 292 LEWREDSRLAFHHVARELHPEYFKNCTSEPYDPR-----LEPWKHFGSQPKTACTNCY 344  
Db 2903 lkwkfdslahrharelhpefykdcxhrrdfprglstgafxews--ggtgsgqpsnaa 2960  
QY 345 CKKCFHC 352  
Db 2961 ykqlifac 2968

Search completed: August 26, 2002, 08:12:30  
Job time: 239 sec

GenCore version 4.5  
Copyright (c) 1993 - 2000 Compugen Ltd.

OM protein - protein search, using sw model

Run on: August 26, 2002, 08:13:22 ; Search time 50.45 Seconds  
(without alignments)  
786.618 Million cell updates/sec

Title: US-09-509-239-21  
Perfect score: 2264  
Sequence: 1 WDPSSHSSNMANTQKSDKI.....QSRGDTGPKETSGHHHHH 413

Scoring table: BLOSUM62  
Gapop 10.0 , Gapext 0.5

Searched: 283138 seqs, 96089334 residues

Total number of hits satisfying chosen parameters: 283138

Minimum DB seq length: 0  
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

Database : PIR.71.\*  
1: pir1.\*  
2: pir2.\*  
3: pir3.\*  
4: pir4.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

## SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	1130	49.9	206	1 ASLJFV	nef protein - huma
2	1115	49.2	205	2 S43467	nef protein - huma
3	1115	49.2	206	2 S03244	nef protein (clone
4	1114	49.2	206	1 ASLJYL	nef protein - huma
5	1108	48.9	206	1 ASLJ12	nef protein - huma
6	1097	48.5	206	2 JC5400	nef protein - huma
7	1096	48.4	206	2 S03245	nef protein (clone
8	1096	48.4	206	2 S33986	nef protein - huma
9	1081.5	47.8	205	1 ASLJH3	nef protein - huma
10	1019	45.0	204	2 S24985	nef protein - huma
11	1019	45.0	218	1 ASLJBR	nef protein - huma
12	1007	44.5	206	2 S25937	nef protein - huma
13	997	44.0	210	1 ASLJ02	nef protein - huma
14	991.5	43.8	214	1 I44001	nef protein - huma
15	987	43.6	182	2 S03247	nef protein (clone
16	987	43.6	204	2 S03246	nef protein - huma
17	911	40.2	212	1 QOLJ2R	nef protein - huma
18	903.5	39.9	207	1 QOLJND	nef protein - huma
19	837	37.0	205	1 B44963	nef protein - huma
20	834.5	36.9	209	2 T01673	nef protein - huma
21	832	36.7	205	1 ASLJIK	nef protein - huma
22	569.5	25.2	364	2 S59931	glycerophosphodies
23	569.5	25.2	364	2 S59932	glycerophosphodies
24	569.5	25.2	364	2 A43576	glycerophosphodies
25	569.5	25.2	364	2 S59934	glycerophosphodies
26	563.5	24.9	364	2 S59933	glycerophosphodies
27	553.5	24.4	364	2 G64086	glycerophosphodies
28	547.5	24.2	364	2 S59935	glycerophosphodies
29	486	21.5	86	2 S33982	trans-activating t

30	478	21.1	95	1 TNLJ12	trans-activating t
31	470	20.8	86	2 A25700	trans-activating t
32	443	19.6	226	2 S46353	nef protein - simi
33	439.5	19.4	97	2 S54385	nef protein - huma
34	437	19.3	101	2 T09446	tat protein - huma
35	434	19.2	101	1 E44001	trans-activating t
36	429.5	19.0	309	2 S07993	nef protein - simi
37	426.5	18.8	259	2 T11567	nef protein - simi
38	421.5	18.6	255	2 S53099	nef protein - huma
39	421	18.6	257	2 S12160	nef protein - huma
40	413	18.2	260	1 ASLJC2	nef protein - huma
41	411.5	18.2	263	2 T11556	nef protein - simi
42	411	18.2	86	2 JC5391	transactivator pro
43	409	18.1	250	2 S54851	nef protein - simi
44	408	18.0	86	2 S54381	tat protein - huma
45	405	17.9	86	1 TNLJ2R	trans-activating t

## ALIGNMENTS

RESULT 1

ASLJFV  
nef protein - human immunodeficiency virus type 1 (isolate LAV-1a)

N:Alternate names: 3'-orf protein; Orf-F protein

C:Species: human immunodeficiency virus type 1, HIV-1

A:Note: host Homo sapiens (man)

C:Date: 17-May-1985 #sequence\_revision 17-May-1985 #text\_change 16-Jul-1999

C:Accession: A04008; S14609

R:Wain-Hobson, S.; Sonigo, P.; Danos, O.; Cole, S.; Allison, M.

Cell 40, 9-17, 1985

A:Title: Nucleotide sequence of the AIDS virus, LAV.

A:Reference number: A90866; MUID:85099333

A:Accession: A04008

A:Molecule type: DNA

A:Residues: 1-206 <WAI>

A:Cross-references: GB:R02013; NID:g326417; PIDN:AAB59752.1; PID:g326425

A:Experimental source: Isolate LAV-1a

R:Ciccarelli, R.B.

submitted to the EMBL Data Library, March 1991

A:Reference number: S14607

A:Accession: S14609

A>Status: preliminary

A:Molecule type: DNA

A:Residues: 1-206 <CIC>

A:Cross-references: EMBL:X58780; NID:g60113; PIDN:CAA1595.1; PID:g60114

C:Genetics:

A:Gene: nef; 3'-orf; Orf-F

C:Superfamily: AIDS nef protein

C:Keywords: AIDS; immunodeficiency; phosphoprotein

Query Match 49.9% ; Score 1130; DB 1; Length 206;  
Best Local Similarity 100.0%; Pred. No. 2.2e-80;  
Matches 206; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 112 MGGKWSKSVVGVPTVRMRRAEPAADGVGAASRDLEKHGAITSNTAATNAACAWLEA 171

Db 1 MGGKWSKSVVGVPTVRMRRAEPAADGVGAASRDLEKHGAITSNTAATNAACAWLEA 60

Qy 172 QEEVEVGFPVTPQVPLRPMYKAAVDLSHFLKEKGLCLHSQRRODLDLWYHTOGY 231

Db 61 QEEVEVGFPVTPQVPLRPMYKAAVDLSHFLKEKGLCLHSQRRODLDLWYHTOGY 120

Qy 232 FPDWQNYTGPQVRYPLTFGWCYKLVVPDPKVEEANKGENTSLHPVSLHGMDDPEREV 291

Db 121 FPDWQNYTGPQVRYPLTFGWCYKLVVPDPKVEEANKGENTSLHPVSLHGMDDPEREV 180

Qy 292 LEWRFDSRLAFHHVARELHPEYFKNC 317

Db 181 LEWRFDSRLAFHHVARELHPEYFKNC 206

QY 232 FPDWQNI TPGGVR YPLTF GWC I KLVF VEPDKVEEANKGENTSLTAPHVSLHGMDDPEREV 291

Query Match 48.9%; Score 1108; DB 1; Length 206;

Matches 200; Conservative 4; Mismatches 2; Indels 0; Gaps 0;

Qy 112 MGGKWSKSSVVGWPTVRMRRAEPAADGVGAASRDLEKHGAITSNTAATNAACAWLEA 171  
 Db 1 MGGKWSKSSVVGWPTVRMRRAEPAADGVGAASRDLEKHGAITSNTAATNAACAWLEA 60

Qy 172 QEEVEGFPVTPQVPLRPMTYKAAVDLSHFLKEKGLGSLHSQRQDILDWIYHTQGY 231  
 Db 61 QEEKVGFPVTPQVPLRPMTYKAAVDLSHFLKEKGLGSLHSQRQDILDWIYHTQGY 120

Qy 232 FPDQNYTPGCVRYPLTFGWCYKLVPEPKVEEANKGENTSLHHPVSLHGMDDPEREV 291  
 Db 121 FPDQNYTPGCVRYPLTFGWCYKLVPEPKVEEANKGENTSLHHPVSLHGMDDPEREV 180

Qy 292 LEWRDLSRLAFHHVARELHPEYFKNC 317  
 Db 181 LEWRDLSRLAFHHVARELHPEYFKNC 206

RESULT 6  
 JC5400  
 nef protein - Human immunodeficiency virus type 1, HIV-1  
 C:Species: Human immunodeficiency virus type 1, HIV-1  
 C:Date: 07-Jul-1997 #sequence\_revision 18-Jul-1997 #text\_change 31-Oct-1997  
 C:Accession: JC5400  
 R:Macreadie, I.G.; Lowe, M.G.; Curtain, C.C.; Hewish, D.; Azad, A.A.  
 Biochem. Biophys. Res. Commun. 232, 707-711, 1997  
 A:Title: Cytotoxicity resulting from addition of HIV-1 Nef N-terminal peptides to yeast  
 A:Reference number: JC5400; MUID:97271389  
 A:Accession: JC5400  
 A:Molecule type: protein  
 A:Residues: 1-206 <MAC>  
 C:Comment: The amino-terminal part possesses membrane-perturbing and fusogenic activities.  
 C:Superfamily: AIDS nef protein  
 F:2-206/Product: nef protein #status predicted <MAT>

Query Match 48.5%; Score 1097; DB 2; Length 206;  
 Best Local Similarity 97.1%; Pred. No. 8.1e-78;  
 Matches 200; Conservative 1; Mismatches 5; Indels 0; Gaps 0;

Qy 112 MGGKWSKSSVVGWPTVRMRRAEPAADGVGAASRDLEKHGAITSNTAATNAACAWLEA 171  
 Db 1 MGGKWSKSSVVGWPTVRMRRAEPAADGVGAASRDLEKHGAITSNTAATNAACAWLEA 60

Qy 172 QEEVEGFPVTPQVPLRPMTYKAAVDLSHFLKEKGLGSLHSQRQDILDWIYHTQGY 231  
 Db 61 QEEVEGFPVTPQVPLRPMTYKAAVDLSHFLKEKGLGSLHSQRQDILDWIYHTQGY 120

Qy 232 FPDQNYTPGCVRYPLTFGWCYKLVPEPKVEEANKGENTSLHHPVSLHGMDDPEREV 291  
 Db 121 FPDQNYTPGCVRYPLTFGWCYKLVPEPKVEEANKGENTSLHHPVSLHGMDDPEREV 180

Qy 292 LEWRDLSRLAFHHVARELHPEYFKNC 317  
 Db 181 LEWRDLSRLAFHHVARELHPEYFKNC 206

RESULT 7  
 S03245  
 nef protein (clone HXB3) - human immunodeficiency virus type 1  
 N:Alternate names: 3'-orf protein  
 C:Species: human immunodeficiency virus type 1, HIV-1  
 C:Date: 28-Feb-1990 #sequence\_revision 28-Feb-1990 #text\_change 20-Sep-1999  
 C:Accession: S03245  
 R:Ratner, L.; Starcich, B.; Josephs, S.F.; Hahn, B.H.; Reddy, E.P.; Livak, K.J.; Pettawa  
 Nucleic Acids Res. 13, 8219-8229, 1985  
 A:Title: Polymorphism of the 3' open reading frame of the virus associated with the acqu  
 A:Reference number: S03244; MUID:86067228  
 A:Accession: S03245  
 A:Molecule type: DNA  
 A:Residues: 1-206 <RAT>

A:Cross-references: EMBL:X03188; NID:961556; PIDN:CAA26947.1; PID:961557  
 C:Genetics:  
 A:Gene: nef; 3'-orf; Orf-F  
 C:Superfamily: AIDS nef protein

Query Match 48.4%; Score 1096; DB 2; Length 206;  
 Best Local Similarity 96.6%; Pred. No. 9.7e-78;  
 Matches 199; Conservative 4; Mismatches 3; Indels 0; Gaps 0;

Qy 112 MGGKWSKSSVVGWPTVRMRRAEPAADGVGAASRDLEKHGAITSNTAATNAACAWLEA 171  
 Db 1 MGGKWSKSSVVGWPTVRMRRAEPAADGVGAASRDLEKHGAITSNTAATNAACAWLEA 60

Qy 172 QEEVEGFPVTPQVPLRPMTYKAAVDLSHFLKEKGLGSLHSQRQDILDWIYHTQGY 231  
 Db 61 QEEKVGFPVTPQVPLRPMTYKAAVDLSHFLKEKGLGSLHSQRQDILDWIYHTQGY 120

Qy 232 FPDQNYTPGCVRYPLTFGWCYKLVPEPKVEEANKGENTSLHHPVSLHGMDDPEREV 291  
 Db 121 FPDQNYTPGCVRYPLTFGWCYKLVPEPKVEEANKGENTSLHHPVSLHGMDDPEREV 180

Qy 292 LEWRDLSRLAFHHVARELHPEYFKNC 317  
 Db 181 LEWRDLSRLAFHHVARELHPEYFKNC 206

RESULT 8  
 S33986  
 nef protein - human immunodeficiency virus type 1  
 C:Species: human immunodeficiency virus type 1, HIV-1  
 C:Date: 19-Mar-1997 #sequence\_revision 19-Mar-1997 #text\_change 20-Sep-1999  
 C:Accession: S33986  
 R:Carlini, F.  
 submitted to the EMBL Data Library, November 1991  
 A:Reference number: S33979  
 A:Accession: S33986  
 A:Status: preliminary  
 A:Molecule type: mRNA  
 A:Residues: 1-206 <CAR>  
 A:Cross-references: EMBL:Z11530; NID:960192; PIDN:CAA77629.1; PID:g60200  
 C:Superfamily: AIDS nef protein

Query Match 48.4%; Score 1096; DB 2; Length 206;  
 Best Local Similarity 96.6%; Pred. No. 9.7e-78;  
 Matches 199; Conservative 3; Mismatches 4; Indels 0; Gaps 0;

Qy 112 MGGKWSKSSVVGWPTVRMRRAEPAADGVGAASRDLEKHGAITSNTAATNAACAWLEA 171  
 Db 1 MGGKWSKSSVVGWPTVRMRRAEPAADGVGAASRDLEKHGAITSNTAATNAACAWLEA 60

Qy 172 QEEVEGFPVTPQVPLRPMTYKAAVDLSHFLKEKGLGSLHSQRQDILDWIYHTQGY 231  
 Db 61 QEEVEGFPVTPQVPLRPMTYKAAVDLSHFLKEKGLGSLHSQRQDILDWIYHTQGY 120

Qy 232 FPDQNYTPGCVRYPLTFGWCYKLVPEPKVEEANKGENTSLHHPVSLHGMDDPEREV 291  
 Db 121 FPDQNYTPGCVRYPLTFGWCYKLVPEPKVEEANKGENTSLHHPVSLHGMDDPEREV 180

Qy 292 LEWRDLSRLAFHHVARELHPEYFKNC 317  
 Db 181 LEWRDLSRLAFHHVARELHPEYFKNC 206

RESULT 9  
 ASLJH3  
 nef protein - human immunodeficiency virus type 1 (isolate HTLV-III, BH10)  
 N:Alternate names: 3'-orf protein; orf-F protein  
 C:Species: human immunodeficiency virus type 1, HIV-1  
 A:Note: host Homo sapiens (man)  
 C:Date: 17-May-1985 #sequence\_revision 17-May-1985 #text\_change 31-Jan-1997  
 C:Accession: A04005

R:Ratner, L.; Haseltine, W.; Patarca, R.; Livak, K.J.; Starcich, B.; Josephs, S.F.; Dora-  
nberger, J.A.; Papas, T.S.; Grayeb, J.; Chang, N.T.; Gallo, R.C.; Wong-Staal, F.  
Nature 313, 277-284, 1985

A:Title: Complete nucleotide sequence of the AIDS virus, HTLV-III.

A:Reference number: A93353; MUID:85111123

A:Accession: A04005

A:Molecule type: DNA

A:Residues: 1-205 <RAT>

C:Genetics:

A:Gene: nef; 3'-orf; orf-F

C:Superfamily: AIDS nef protein

C:Keywords: AIDS; immunodeficiency

Query Match 47.8%; Score 1081.5; DB 1; Length 205;  
Best Local Similarity 96.1%; Pred. No. 1.3e-76;  
Matches 198; Conservative 4; Mismatches 3; Indels 1; Gaps 1;

Qy 112 MGGKWSKSSVVGWPTVRRMRRAEPAADGCGAASRDLEKKGAISSNTAATNAACAWLEA 171

Db 1 MGGKWSKSSVVGWPAVRRMRRAEPAADGCGAASRDLEKKGAISSNTAANNACAWLEA 60

Qy 172 QEEVEGFPVTPQVPLRPMYKAAVDLSHFLKEKGGLEGLHSORQDILDLYIHTQGY 231

Db 61 QEEVEGFPVTPQVPLRPMYKAAVDLSHFLKEKGGLEGLHSORQDILDLYIHTQGY 120

Qy 232 FPDQNTPGGVRYPLTFGCYKLVPEPKVEANKGENTSLHPVSLHGMDDPREV 291

Db 121 FPD-QNTPGGVRYPLTFGCYKLVPEPKVEANKGENTSLHPVSLHGMDDPREV 179

Qy 292 LEWRDLSRLAFHHVARELHPEYFKNC 317

Db 180 LEWRDLSRLAFHHVARELHPEYFKNC 205

RESULT 10

S24985

nef protein - human immunodeficiency virus type 1

C:Species: human immunodeficiency virus type 1, HIV-1

C:Date: 20-Feb-1995 #sequence\_revision 20-Feb-1995 #text\_change 20-Sep-1999

C:Accession: S24985

R:Harris, M.; Hislop, S.; Patsillacos, P.; Neil, J.C.

submitted to the EMBL Data Library, November 1991

A:Description: In vivo derived HIV-1 NEF genes are heterogeneous and lack detectable nuc

A:Reference number: S24985

A:Accession: S24985

A>Status: preliminary

A:Molecule type: DNA

A:Residues: 1-204 <HAR>

A:Cross-references: EMBL:X63044; NID:g60123; PIDN:CAA44770.1; PID:g60124

C:Superfamily: AIDS nef protein

Query Match 45.0%; Score 1019; DB 2; Length 204;  
Best Local Similarity 89.8%; Pred. No. 9.1e-72;  
Matches 185; Conservative 10; Mismatches 9; Indels 2; Gaps 1;

Qy 112 MGGKWSKSSVVGWPTVRRMRRAEPAADGCGAASRDLEKKGAISSNTAATNAACAWLEA 171

Db 1 MGGKWSK--LGWPTVRRMRRAEPAADGCGAASRDLEKKGAISSNTAATNAACAWLEA 58

Qy 172 QEEVEGFPVTPQVPLRPMYKAAVDLSHFLKEKGGLEGLHSORQDILDLYIHTQGY 231

Db 59 QEEVEGFPVTPQVPLRPMYKAAVDLSHFLKEKGGLEGLHSORQDILDLYIHTQGY 118

Qy 232 FPDQNTPGGVRYPLTFGCYKLVPEPKVEANKGENTSLHPVSLHGMDDPREV 291

Db 119 FPDQNTPGGVRYPLTFGCYKLVPEPKVEANKGENTSLHPVSLHGMDDPREV 178

Qy 292 LEWRDLSRLAFHHVARELHPEYFKNC 317

Db 179 LEWRDLSRLAFHHVARELHPEYFKNC 204

RESULT 11

ASLJBR

nef protein - human immunodeficiency virus type 1 (isolate BR)

N:Alternate names: 3'-orf protein; orf-F protein

C:Species: human immunodeficiency virus type 1, HIV-1

A>Note: host Homo sapiens (man)

C:Date: 31-Dec-1989 #sequence\_revision 31-Dec-1989 #text\_change 16-Jul-1999

C:Accession: D31667; S21993; S21995; S21997; S21999; S21991

R:Anand, R.; Thayer, R.; Srinivasan, A.; Nayyar, S.; Gardner, M.; Luciw, P.; Dandekar

Virology 168, 79-89, 1989

A:Title: Biological and molecular characterization of human immunodeficiency virus (I

A:Reference number: A94389; MUID:89085613

A:Accession: D31667

A:Molecule type: DNA

A:Residues: 1-218 <ANA>

A:Cross-references: GB:M21098; NID:g326426; PIDN:AAA44222.1; PID:g326431

R:Steuler, H.; Storch-Hagenlocher, B.; Wildemann, B.; Hacke, W.

submitted to the EMBL Data Library, July 1991

A:Description: Distinct populations of HIV-1 in blood and cerebrospinal fluid as det

A:Reference number: S21990

A:Accession: S21993

A:Molecule type: DNA

A:Residues: 1-7 <STE>

A:Cross-references: EMBL:X61358; NID:g60177; PIDN:CAA43629.1; PID:g584027; EMBL:X613

A:Accession: S21995

A:Molecule type: DNA

A:Residues: 1-7 <ST2>

A:Cross-references: EMBL:X61355; NID:g60179; PIDN:CAA43623.1; PID:g584028; EMBL:X613

A:Accession: S21997

A:Molecule type: DNA

A:Residues: 1-7 <ST3>

A:Cross-references: EMBL:X61356; NID:g60181; PIDN:CAA43625.1; PID:g1129140; EMBL:X61

A:Accession: S21999

A:Molecule type: DNA

A:Residues: 1-7 <ST4>

A:Cross-references: EMBL:X61359; NID:g60182; PIDN:CAA43631.1; PID:g584030

C:Genetics:

A:Gene: nef; 3'-orf; orf-F

C:Superfamily: AIDS nef protein

C:Keywords: AIDS; immunodeficiency

Query Match 45.0%; Score 1019; DB 1; Length 218;

Best Local Similarity 84.1%; Pred. No. 9.9e-72;

Matches 185; Conservative 12; Mismatches 7; Indels 16; Gaps 2;

Qy 112 MGGKWSKSSVVGWPTVRRMR-----RAEPAADGCGAASRDLEKKGAISS 157

Db 1 MGGKWSK--MAGWSTVRRMRRAEPAEPAADGCGAASRDLEKKGAISS 58

Qy 158 NTAATNAACAWLEAQEEVEGFPVTPQVPLRPMYKAAVDLSHFLKEKGGLEGLHSORR 217

Db 59 NTAATNADCAWLEAQEEVEGFPVTPQVPLRPMYKAAVDLSHFLKEKGGLEGLHSORR 118

Qy 218 QDILDLYIHTQGYPPDQNTPGGVRYPLTFGCYKLVPEPKVEANKGENTSLH 277

Db 119 QDILDLYIHTQGYPPDQNTPGGVRYPLTFGCYKLVPEPKVEANKGENTSLH 178

Qy 278 PVSLSHGMDPREVLEWRDLSRLAFHHVARELHPEYFKNC 317

Db 179 PMSQHGMDPREVLEWRDLSRLAFHHVARELHPEYFKNC 218

RESULT 12

S25937

nef protein - human immunodeficiency virus type 1

C:Species: human immunodeficiency virus type 1, HIV-1

C:Date: 25-Feb-1994 #sequence\_revision 01-Dec-1995 #text\_change 20-Sep-1999

C:Accession: S25937

R:Guo, H.G.; Chermann, J.C.; Waters, D.; Hall, L.; Louie, A.; Gallo, R.C.; Streicher

Nature 349, 745-746, 1991

A:Title: Sequence analysis of original HIV-1.

A:Reference number: S25937; MUID:91156044  
A:Accession: S25937  
A>Status: preliminary; nucleic acid sequence not shown; translation not shown  
A:Molecule type: nucleic acid  
A:Residues: 1-206 <GUO>  
A:Cross-references: EMBL:X57465; NID:g60217; PIDN:CRAA0702.1; PID:g60218  
A>Note: the nucleotide sequence was submitted to the EMBL Data Library, March 1991  
C:Superfamily: AIDS nef protein

Query Match 44.5%; Score 1007; DB 2; Length 206;  
Best Local Similarity 85.9%; Pred. No. 7.9e-71;  
Matches 177; Conservative 18; Mismatches 11; Indels 0; Gaps 0;

QY 112 MGCKWSKSSVGVGPTVRRMRRAEPAADGVGAASRDLEKHGAITSSNTAATNAACAWLEA 171  
Db 1 MGCKWSKSSVGVGPTVRRMRRAEPAADGVGAASRDLEKHGAITSSNTAATNAACAWLEA 60

QY 172 QEEEEVGFPTVQVPLRPMYTKAAVDLSHFLKKEGGLIHSORQDILDLMWYHTQGY 231  
Db 61 QEEKVGFPVQVPLRPMYTKGALDLSHFLKKGGLIYSQKQDILDLMWYHTQGY 120

QY 232 FPDQNYTGPQVRYPLTFGWCYKLVPEPKVEEANKGENTSLHPVSLHGMDDPEREV 291  
Db 121 FPDQNYTGPQVRYPLTFGWCYKLVPEPKVEEANKGENTSLHPVSLHGMDDPEREV 180

QY 292 LEWRFDRLAFHHVARELHPEYFKNC 317  
Db 181 LEWRFDRLAFHHVARELHPEYFKNC 206

RESULT 13  
ASLJ02  
nef protein - human immunodeficiency virus type 1 (isolate ARV-2)  
N:Alternate names: 3'-orf protein; orf-F protein  
C:Species: human immunodeficiency virus type 1, HIV-1  
A>Note: host Homo sapiens (man)  
C>Date: 17-May-1985 #sequence\_revision 17-May-1985 #text\_change 16-Jul-1999  
C:Accession: A04009  
R:Sanchez-Pescador, R.; Power, M.D.; Barr, P.J.; Steimer, K.S.; Stemplen, M.M.; Brown-SH  
Science 227, 484-492, 1985  
A:Title: Nucleotide sequence and expression of an AIDS-associated retrovirus (ARV-2).  
A:Reference number: A04003; MUID:85090453  
A:Accession: A04009  
A:Molecule type: DNA  
A:Residues: 1-210 <SAN>  
A:Cross-references: GB:K02007; NID:g328658; PIDN:AAB59883.1; PID:g328667  
C:Genetics:  
A:Gene: nef; 3'-orf; orf-F  
C:Superfamily: AIDS nef protein  
C:Keywords: AIDS; immunodeficiency

Query Match 44.0%; Score 997; DB 1; Length 210;  
Best Local Similarity 85.2%; Pred. No. 4.8e-70;  
Matches 179; Conservative 16; Mismatches 11; Indels 4; Gaps 1;

QY 112 MGCKWSKSSVGVGPTVRRMRRAEPAADGVGAASRDLEKHGAITSSNTAATNAACAWLEA 167  
Db 1 MGCKWSKSSVGVGPTVRRMRRAEPAADGVGAASRDLEKHGAITSSNTAATNAACAWLEA 60

QY 168 WLEAQEEVEGFPVTPQVPLRPMYTKAAVDLSHFLKKEGGLIHSORQDILDLMWYH 227  
Db 61 WLEAQEEVEGFPVTPQVPLRPMYTKAAVDLSHFLKKEGGLIHSORQDILDLMWYH 120

QY 228 TQGYFPDQNYTGPQVRYPLTFGWCYKLVPEPKVEEANKGENTSLHPVSLHGMDDP 287  
Db 121 TQGYFPDQNYTGPQVRYPLTFGWCYKLVPEPKVEEANKGENTSLHPVSLHGMDDP 180

QY 288 EREVLEWRFDRLAFHHVARELHPEYFKNC 317  
Db 181 EREVLEWRFDRLAFHHVARELHPEYFKNC 210

RESULT 14  
I44001  
nef protein - human immunodeficiency virus type 1 (strain YU-2)  
N:Alternate names: 3'-orf protein; orf-F protein  
C:Species: human immunodeficiency virus type 1, HIV-1  
A>Note: host Homo sapiens (man)  
C>Date: 31-Dec-1993 #sequence\_revision 31-Dec-1993 #text\_change 23-Feb-1997  
C:Accession: I44001  
R:Li, Y.; Hul, H.; Burgess, C.J.; Price, R.W.; Sharp, P.M.; Hahn, B.H.; Shaw, G.M.  
J. Virol. 66, 6587-6600, 1992  
A:Title: Complete nucleotide sequence, genome organization, and biological properties  
A:Reference number: A44001; MUID:93021387  
A:Accession: I44001  
A:Molecule type: DNA  
A:Residues: 1-214 <LIY>  
A:Cross-references: GB:M93258  
C:Genetics:  
A:Gene: nef; 3'-orf; orf-F  
C:Superfamily: AIDS nef protein  
C:Keywords: AIDS; immunodeficiency

Query Match 43.8%; Score 991.5; DB 1; Length 214;  
Best Local Similarity 84.2%; Pred. No. 1.3e-69;  
Matches 181; Conservative 12; Mismatches 11; Indels 11; Gaps 2;

QY 112 MGCKWSKSSVGVGPTVRRMRRAEPAADGVGAASRDLEKHGAITSSNTAA 161  
Db 1 MGCKWSKSSVGVGPTVRRMRRAEPAADGVGAASRDLEKHGAITSSNTAA 60

QY 162 TNACAWLEAQEEVEGFPVTPQVPLRPMYTKAAVDLSHFLKKEGGLIHSORQDIL 221  
Db 61 TNACAWLEAQEEVEGFPVTPQVPLRPMYTKAAVDLSHFLKKEGGLIHSORQDIL 120

QY 222 DLWYHTQGYFPDQNYTGPQVRYPLTFGWCYKLVPEPKVEEANKGENTSLHPVSL 281  
Db 121 DLWYHTQGYFPDQNYTGPQVRYPLTFGWCYKLVPEPKVEEANKGENTSLHPVSL 179

QY 282 HGMDDPEREVRFDRLAFHHVARELHPEYFKN 316  
Db 180 HGMDDPEREVRFDRLAFHHVARELHPEYFKN 214

RESULT 15  
S03247  
nef protein (clone C15) - human immunodeficiency virus type 1 (fragment)  
N:Alternate names: 3'-orf protein  
C:Species: human immunodeficiency virus type 1, HIV-1  
C>Date: 28-Feb-1990 #sequence\_revision 28-Feb-1990 #text\_change 20-Jun-2000  
C:Accession: S03247  
R:Ratner, L.; Starcich, B.; Josephs, S.F.; Hahn, B.H.; Reddy, E.P.; Livak, K.J.; Pett  
Nucleic Acids Res. 13, 8219-8229, 1985  
A:Title: Polymorphism of the 3' open reading frame of the virus associated with the a  
A:Reference number: S03244; MUID:86067228  
A:Accession: S03247  
A:Molecule type: mRNA  
A:Residues: 1-182 <RAT>  
A:Cross-references: EMBL:X03189; NID:g61552; PIDN:CAA26948.1; PID:g1335562  
C:Genetics:  
A:Gene: nef; 3'-orf; orf-F  
C:Superfamily: AIDS nef protein

Query Match 43.6%; Score 987; DB 2; Length 182;  
Best Local Similarity 98.4%; Pred. No. 2.4e-69;  
Matches 179; Conservative 1; Mismatches 2; Indels 0; Gaps 0;

QY 136 PAADGVGAASRDLEKHGAITSSNTAATNAACAWLEAQEEVEGFPVTPQVPLRPMYTKAA 195  
Db 1 PAADGVGAASRDLEKHGAITSSNTAATNAACAWLEAQEEVEGFPVTPQVPLRPMYTKAA 60

QY 196 VDLSHFLKKEGGLIHSORQDILDLMWYHTQGYFPDQNYTGPQVRYPLTFGWCYK 255

Db 61 VDLSHFLKEKGLGLHSQRQDLDLWYHTQGYFPDMQNYTPGCVRIPLTFGWCYK 120

Qy 256 LVPVEPKVEANKGENTSLLRHPSVSLHGMDDPEREVLWFRDLSRLAFHHVARELHPYFK 315

Db 121 LVPVEPKVEANKGENTSLLRHPSVSLHGMDDPEREVLWFRDLSRLAFHHVARELHPYFK 180

Qy 316 NC 317

Db 181 NC 182

## RESULT 16

S03246

nef protein (clone HAT3) - human immunodeficiency virus type 1

N:Alternate names: 3'-orf protein

C:Species: human immunodeficiency virus type 1, HIV-1

C:Date: 28-Feb-1990 #sequence\_revision 28-Feb-1990 #text\_change 20-Sep-1999

A:Accession: S03246

R:Ratner, L.; Starcich, B.; Josephs, S.F.; Hahn, B.H.; Reddy, E.P.; Livak, K.J.; Pettew

Nucleic Acids Res. 13, 8219-8229, 1985

A:Title: Polymorphism of the 3' open reading frame of the virus associated with the acqu

A:Reference number: S03244; MUID:86067228

A:Accession: S03246

A:Molecule type: DNA

A:Residues: 1-204 &lt;RAT&gt;

A:Cross-references: EMBL:X03190; NID:g61550; PIDN:CAA26949.1; PID:g61551

A:Note: the authors translated the codon AGT for residue 11 as Gly

C:Genetics:

A:Gene: nef; 3'-orf; orf-F

C:Superfamily: AIDS nef protein

Query Match 43.6%; Score 987; DB 2; Length 204;  
Best Local Similarity 84.5%; Pred. No. 2.8e-69;  
Matches 174; Conservative 16; Mismatches 14; Indels 2; Gaps 1;

Qy 112 MGGKWSKSSVVGWPTVRERMRRAEPAADGVGAASRDLEKKGALTSNTAATNAACAWLEA 171

Db 1 MGGKWSKSSVVGWPTVRERMRRAEPAADGVGAASRDLEKKGALTSNTAATNAACAWLEA 58

Qy 172 QEEVEGFPVTPQVPLRPMTYKAAVDLSHFLKEKGLGLHSQRQDLDLWYHTQGY 231

Db 59 QEEVEGFPVTPQVPLRPMTYKAAVDLSHFLKEKGLGLHSQRQDLDLWYHTQGY 118

Qy 232 FPDQNYTPGCVRIPLTFGWCYKLVVPVEPKVEANKGENTSLLRHPSVSLHGMDDPERE 291

Db 119 FPDQNYTPGCVRIPLTFGWCYKLVVPVEPKVEANKGENTSLLRHPSVSLHGMDDPERE 178

Qy 292 LEWRFDSRLAFHHVARELHPYFKNC 317

Db 179 LVWKFDLSRLAFHHVARELHPYFKNC 204

## RESULT 17

Q0LJ2R

nef protein - human immunodeficiency virus 2r-6

N:Alternate names: 3'-orf protein; orf-F protein

C:Species: human immunodeficiency virus 2r-6

C:Date: 30-Sep-1987 #sequence\_revision 30-Sep-1987 #text\_change 16-Jul-1999

A:Accession: F26192

R:Srinivasan, A.; Anand, R.; York, D.; Ranganathan, P.; Feorino, P.; Schochetman, G.; Cu

Gene 52, 71-82, 1987

A:Title: Molecular characterization of human immunodeficiency virus from Zaire: nucleot

A:Reference number: A26192; MUID:87248097

A:Accession: F26192

A:Molecule type: DNA

A:Residues: 1-212 &lt;SRI&gt;

A:Cross-references: GB:K03458; GB:M16322; NID:g29398; PIDN:AAA45381.1; PID:g29404

C:Genetics:

A:Gene: nef; 3'-orf; orf-F

C:Superfamily: AIDS nef protein

C:Keywords: AIDS; immunodeficiency

Query Match 40.2%; Score 911; DB 1; Length 212;

Best Local Similarity 77.4%; Pred. No. 2.3e-63;

Matches 164; Conservative 23; Mismatches 19; Indels 6; Gaps 2;

Qy 112 MGGKWSKSSVVGWPTVRERMRRAEPAADGVGAASRDLEKKGALTSNTAATNAAC 166

Db 1 MGGKWSKSSVVGWPTVRERMRRAEPAADGVGAASRDLEKKGALTSNTAATNAAC 60

Qy 167 AWLEAQEE-EEVGFPTQVPLRPMTYKAAVDLSHFLKEKGLGLHSQRQDLDLWY 225

Db 61 AWLEAQEESEVGFPTQVPLRPMTYKAAVDLSHFLKEKGLGLHSQRQDLDLWY 120

Qy 226 YHTQGYFPDMQNYTPGCVRIPLTFGWCYKLVVPVEPKVEANKGENTSLLRHPSVSLHGM 285

Db 121 YHTQGYFPDMQNYTPGCVRIPLTFGWCYKLVVPVEPKVEANKGENTSLLRHPSVSLHGM 180

Qy 286 DPEREVLWFRDLSRLAFHHVARELHPYFKNC 317

Db 181 DPEREVLWFRDLSRLAFHHVARELHPYFKNC 212

## RESULT 18

Q0LJND

nef protein - human immunodeficiency virus type 1 (isolate NDK)

N:Alternate names: 3'-orf protein; orf-F protein

C:Species: human immunodeficiency virus type 1, HIV-1

A:Note: host Homo sapiens (man)

C:Date: 30-Sep-1991 #sequence\_revision 30-Sep-1991 #text\_change 16-Jul-1999

A:Accession: JQ0068

R:Spire, B.; Sire, J.; Zachar, V.; Rey, F.; Barre-Sinoussi, F.; Galibert, F.; Hampe,

Gene 81, 275-284, 1989

A:Title: Nucleotide sequence of HIV1-NDK: a highly cytopathic strain of the human im

A:Reference number: JQ0065; MUID:90034200

A:Accession: JQ0068

A:Molecule type: DNA

A:Residues: 1-207 &lt;SRI&gt;

A:Cross-references: GB:M27323; NID:g328154; PIDN:AAA44874.1; PID:g328163

C:Genetics:

A:Gene: nef

C:Superfamily: AIDS nef protein

C:Keywords: AIDS; immunodeficiency

Query Match 39.9%; Score 903.5; DB 1; Length 207;

Best Local Similarity 77.3%; Pred. No. 8.6e-63;

Matches 160; Conservative 25; Mismatches 21; Indels 1; Gaps 1;

Qy 112 MGGKWSKSSVVGWPTVRERMRRAEPAADGVGAASRDLEKKGALTSNTAATNAACAWLEA 171

Db 1 MGGKWSKSSVVGWPTVRERMRRAEPAADGVGAASRDLEKKGALTSNTAATNAACAWLEA 60

Qy 172 QEE-EEVGFPTQVPLRPMTYKAAVDLSHFLKEKGLGLHSQRQDLDLWYHTQGY 230

Db 61 QEESEVGFPTQVPLRPMTYKAAVDLSHFLKEKGLGLHSQRQDLDLWYHTQGY 120

Qy 231 YFPDMQNYTPGCVRIPLTFGWCYKLVVPVEPKVEANKGENTSLLRHPSVSLHGMDDPERE 290

Db 121 YFPDMQNYTPGCVRIPLTFGWCYKLVVPVEPKVEANKGENTSLLRHPSVSLHGMDDPERE 180

Qy 291 VLEWRFDSRLAFHHVARELHPYFKNC 317

Db 181 VLEWRFDSRLAFHHVARELHPYFKNC 207

## RESULT 19

B44963

nef protein - human immunodeficiency virus type 1 (isolate 2321)

N:Alternate names: 3'-orf protein; orf-F protein

C:Species: human immunodeficiency virus type 1, HIV-1

A:Note: host Homo sapiens (man)

C:Date: 17-Feb-1994 #sequence\_revision 17-Feb-1994 #text\_change 28-May-1999



C:Accession: B44963  
R:Srinivasan, A.; York, D.; Butler Jr., D.; Jannoun-Nasr, R.; Getchell, J.; McCormick, J.  
AIDS Res. Hum. Retroviruses 5, 121-129, 1989  
A:Title: Molecular characterization of HIV-1 isolated from a serum collected in 1976: nu  
A:Reference number: A44963; MUID:89228766  
A:Accession: B44963  
A:Molecule type: DNA  
A:Residues: 1-205 <SRI>  
A:Cross-references: GB:M15896; NID:g329392; PIDN:AAB53951.1; PID:g329397  
C:Genetics:  
A:Gene: nef; 3'-orf; orf-F  
C:Superfamily: AIDS nef protein  
C:Keywords: AIDS; Immunodeficiency

Query Match 37.0%; Score 837; DB 1; Length 205;  
Best Local Similarity 73.2%; Pred. No. 1.2e-57;  
Matches 153; Conservative 23; Mismatches 25; Indels 8; Gaps 3;

QY 112 MGGKWSKSSVVGWPTVRMR---AEPADGVGAASRDLEKHGAITSSNTAATNAACAW 168  
DB 1 MGNKSK-----GWPVREIRQTPTAPPAAEGVGAASQDLAKHGAISSNTATNPDCAW 56

QY 169 LEAQEE-EEVGFPTQVPLRPMYKAAVDLSHFLKKGGLGLIHSQRQDILDLWIYH 227  
DB 57 LEAQEESEEVGFVPRQVPLRPMTEKGFADLSFLKKGGLGLIYSKKRQEIIDLWVYH 116

QY 228 TQGYFPDQWNTPGCVRYPLTFGCWYKLVPEPDKVEANKGENTSLHPVSLHGMDP 287  
DB 117 TQGYFPDQWNTPGCVRYPLTFGCWYKLVPEPDKVEANKGENTSLHPVSLHGMDP 287

QY 288 EREVLEWRFDSRLAFHHVARELHPEYFKN 316  
DB 177 EREVLEWRFDSRLAFHHVARELHPEYFKN 205

RESULT 20  
T01673  
nef protein - human immunodeficiency virus type 1  
C:Species: human immunodeficiency virus type 1, HIV-1  
C:Date: 19-Feb-1999 #sequence\_revision 19-Feb-1999 #text\_change 21-Jul-2000  
C:Accession: T01673  
R:Alizon, M.; Wain-Hobson, S.; Gluckman, J.C.; Sonigo, P.  
Cell 46, 63-74, 1986  
A:Title: Genetic variability of the AIDS virus: Nucleotide sequence analysis of two isol  
A:Reference number: Z14389; MUID:86245056  
A:Accession: T01673  
A>Status: preliminary; translated from GB/EMBL/DBDJ  
A:Molecule type: mRNA  
A:Residues: 1-209 <ALI>  
A:Cross-references: EMBL:K03456; NID:g60228; PIDN:CAA28017.1; PID:g60235  
C:Superfamily: AIDS nef protein

Query Match 36.9%; Score 834.5; DB 2; Length 209;  
Best Local Similarity 70.1%; Pred. No. 2e-57;  
Matches 148; Conservative 26; Mismatches 30; Indels 7; Gaps 2;

QY 112 MGGKWSKSSVVGWPTVRMRRAEPADGVG-----AASRDLEKHGAITSSNTAATNAAC 166  
DB 1 MGGKWSKSSVVGWPKIRIRIRPTTETGTGAVSQDADKCGAASSSPAANNASC 60

QY 167 AWLEAQEESEEVGFPTQVPLRPMYKAAVDLSHFLKKGGLGLIHSQRQDILDLWIY 226  
DB 61 E--PPSEEEVGFVPRQVPLRPMYKGFADLSHFLKKGGLGLIHSQRQDILDLWIY 118

QY 227 HTQGYFPDQWNTPGCVRYPLTFGCWYKLVPEPDKVEANKGENTSLHPVSLHGMD 286  
DB 119 HTQGYFPDQWNTPGVIRFPLTFGCWYKLVPMSPPEVEANEENCLLHPISQHGMD 178

QY 287 PEREVLEWRFDSRLAFHHVARELHPEYFKN 317  
DB 179 AEREVLKWKFDSSLALRHRAHQHPEYFKN 209

RESULT 21  
ASLJIK  
nef protein - simian immunodeficiency virus SIVcpz  
N:Alternate names: 3'-orf protein; orf-F protein  
C:Species: simian immunodeficiency virus SIVcpz  
A:Note: host Pan troglodytes (chimpanzee)  
C:Date: 30-Sep-1991 #sequence\_revision 30-Sep-1991 #text\_change 16-Jul-1999  
C:Accession: S09991  
R:Huet, T.; Cheynier, R.; Meyerhans, A.; Roelants, G.; Wain-Hobson, S.  
Nature 345, 356-359, 1990  
A:Title: Genetic organization of a chimpanzee lentivirus related to HIV-1.  
A:Reference number: S09983; MUID:90259077  
A:Accession: S09991  
A>Status: nucleic acid sequence not shown; translation not shown  
A:Molecule type: DNA  
A:Residues: 1-205 <HUE>  
A:Cross-references: EMBL:X52154; NID:g58866; PIDN:CAA36408.1; PID:g58877  
C:Genetics:  
A:Gene: nef  
C:Superfamily: AIDS nef protein  
C:Keywords: AIDS; Immunodeficiency

Query Match 36.7%; Score 832; DB 1; Length 205;  
Best Local Similarity 70.2%; Pred. No. 3e-57;  
Matches 144; Conservative 27; Mismatches 34; Indels 0; Gaps 0;

QY 112 MGGKWSKSSVVGWPTVRMRRAEPADGVGAASRDLEKHGAITSSNTAATNAACAWLEA 171  
DB 1 MGTGWSKSSVVGWPEVRRIRREAPTAAGVGEVSKOLERHGAITSRTPTNQTALWEE 60

QY 172 QEEEEVGFPTQVPLRPMYKAAVDLSHFLKKGGLGLIHSQRQDILDLWIYHTQGY 231  
DB 61 MDNEEVGFVPRQVPLRPMYKAAFDLSHFLKKGGLGLVSRRRQEIIDLWVYHTQGF 120

QY 232 FPDQWNTPGCVRYPLTFGCWYKLVPEPDKVEANKGENTSLHPVSLHGMDPEREV 291  
DB 121 FPDQWNTPGCVRYPLTFGCWYKLVPEPDKVEANKGENTSLHPVSLHGMDPEREV 180

QY 292 LEWRFDSRLAFHHVARELHPEYFKN 316  
DB 181 LEWRFDSRLAFHHVARELHPEYFKN 205

RESULT 22  
S59931  
glycerophosphodiester phosphodiesterase (EC 3.1.4.46) - Haemophilus influenzae (strai  
N:Alternate names: Igd-binding protein; protein D  
C:Species: Haemophilus influenzae  
A:Variety: strain 3639  
C:Date: 15-Feb-1996 #sequence\_revision 01-Mar-1996 #text\_change 08-Oct-1999  
C:Accession: S59931; S47333  
R:Song, X.M.; Forsgren, A.; Jansson, H.  
Infect. Immun. 63, 696-699, 1995  
A:Title: The gene encoding protein D (hpd) is highly conserved among Haemophilus Infl  
A:Reference number: S59931; MUID:95122210  
A:Accession: S59931  
A>Status: preliminary  
A:Molecule type: DNA  
A:Residues: 1-364 <SON>  
A:Cross-references: EMBL:Z35556; NID:g525213; PIDN:CAA84715.1; PID:g525214  
A:Experimental source: strain 3639  
C:Keywords: lipoprotein; phosphoric diester hydrolase; surface antigen

Query Match 25.2%; Score 569.5; DB 2; Length 364;  
Best Local Similarity 46.9%; Pred. No. 1.4e-36;  
Matches 144; Conservative 19; Mismatches 71; Indels 73; Gaps 11;

QY 4 SSHSSNNANTQMSDKIIIAHRCAGSYLPHTLESKALAFQAQADYLEQDLAMTKDGLV 63  
|||||



F:1-18/Domain: signal sequence #status predicted <SIG>  
F:19-364/Product: protein D #status predicted <MAT>

Query Match 25.2%; Score 569.5; DB 2: Length 364;  
Best Local Similarity 46.9%; Pred. No. 1.4e-36;  
Matches 144; Conservative 19; Mismatches 71; Indels 73; Gaps 11;

QY 4 SSHSSNNANTOMKSKDIIIAHRCASCYLPEHTLESKALAFQAQADYLEODLMTKDGRLV 63  
|||||  
DB 20 SSHSSNNANTOMKSKDIIIAHRCASCYLPEHTLESKALAFQAQADYLEODLMTKDGRLV 79  
|||||  
QY 64 VIHDFLDGLTDVAKKFPFHRKDGYYVIDFTLKEIQSLEMTENFETMGKK----- 115  
|||||  
DB 80 VIHDFLDGLTDVAKKFPFHRKDGYYVIDFTLKEIQSLEMTENFETMGKKQAQVYVNR 139  
|||||  
QY 116 -----W-----SKSSVVG-WPTVVRMRRAEPAADGVGAASRDLEK 150  
|||||  
DB 140 FPLWQSHFRHTFDEIEFIQGLEKSTGRKVGIYPEIKAPWFHONGKDIAAETLKVLLK 199  
|||||  
QY 151 HGAITSNTAATNAACAWLEAQAEEVEGFPVTPQVPLRPMTYKAAVDLSHFLKEKGGLEG 210  
|||||  
DB 200 YGYDKKTD-----MVLQTFDFNELKRIKTELLPQMGMDLK-LVQLIAYTDWK----- 246  
|||||  
QY 211 LIHSORRQDILDLIYHTOGYFPDQW-----NYTPGQVRYPLTFGWCYKLVVPE 260  
|||||  
DB 247 ---ETOEKDPKGYVW---NINYDMWFKPGMAEVVYKADGVGP-----GW-YMLVNKE 292  
|||||  
QY 261 ---PDKV 264  
|||  
DB 293 ESKPDNI 299

RESULT 26  
S59933  
glycerophosphodiester phosphodiesterase (EC 3.1.4.46) - Haemophilus influenzae (strain 6  
N:Alternate names: immunoglobulin D-binding protein; protein D  
C:Species: Haemophilus influenzae  
A:Variety: strain 6-7626  
C:Date: 15-Feb-1996 #sequence\_revision 01-Mar-1996 #text\_change 08-Oct-1999  
C:Accession: S59933; S47335  
R:Song, X.M.; Forsgren, A.; Janson, H.  
Infect. Immun. 63, 696-699, 1995  
A:Title: The gene encoding protein D (hpd) is highly conserved among Haemophilus influen  
A:Reference number: S59931; MUID:95122210  
A:Accession: S59933  
A:Molecule type: DNA  
A:Residues: 1-364 <TIGR>  
A:Cross-references: EMBL:235658; NID:9525217; PIDN:CAA84717.1; PID:9525218  
A:Experimental source: strain 6-7626  
C:Keywords: lipoprotein; phosphoric diester hydrolase; surface antigen

Query Match 24.9%; Score 563.5; DB 2: Length 364;  
Best Local Similarity 46.6%; Pred. No. 4.2e-36;  
Matches 143; Conservative 19; Mismatches 72; Indels 73; Gaps 11;

QY 4 SSHSSNNANTOMKSKDIIIAHRCASCYLPEHTLESKALAFQAQADYLEODLMTKDGRLV 63  
|||||  
DB 20 SSHSSNNANTOMKSKDIIIAHRCASCYLPEHTLESKALAFQAQADYLEODLMTKDGRLV 79  
|||||  
QY 64 VIHDFLDGLTDVAKKFPFHRKDGYYVIDFTLKEIQSLEMTENFETMGKK----- 115  
|||||  
DB 80 VIHDFLDGLTDVAKKFPFHRKDGYYVIDFTLKEIQSLEMTENFETMGKKQAQVYVNR 139  
|||||  
QY 116 -----W-----SKSSVVG-WPTVVRMRRAEPAADGVGAASRDLEK 150  
|||||  
DB 140 FPLWQSHFRHTFDEIEFIQGLEKSTGRKVGIYPEIKAPWFHONGKDIAAETLKVLLK 199  
|||||  
QY 151 HGAITSNTAATNAACAWLEAQAEEVEGFPVTPQVPLRPMTYKAAVDLSHFLKEKGGLEG 210  
|||||  
DB 200 YGYDKKTD-----MVLQTFDFNELKRIKTELLPQMGMDLK-LVQLIAYTDWK----- 246  
|||||

QY 211 LIHSORRQDILDLIYHTOGYFPDQW-----NYTPGQVRYPLTFGWCYKLVVPE 260  
|||||  
DB 247 ---ETOEKDSKGYVW---NINYDMWFKPGMAEVVYKADGVGP-----GW-YMLVNKE 292  
|||||  
QY 261 ---PDKV 264  
|||  
DB 293 ESKPDNI 299

RESULT 27  
G64086  
glycerophosphodiester phosphodiesterase (EC 3.1.4.46) - Haemophilus influenzae (strai  
N:Alternate names: Igp-binding protein; protein D  
C:Species: Haemophilus influenzae  
C:Date: 18-Aug-1995 #sequence\_revision 18-Aug-1995 #text\_change 21-Jul-2000  
C:Accession: G64086  
R:Feilschmann, R.D.; Adams, M.D.; White, O.; Clayton, R.A.; Kirkness, E.F.; Kerlavage  
; Gocayne, J.D.; Scott, J.; Shirley, R.; Liu, L.I.; Glodek, A.; Kelley, J.M.; Weidman  
; D.M.; Brandon, R.C.; Fine, L.D.; Fritchman, J.L.; Fuhmann, J.L.; Geoghagen, N.S.M.  
Science 269, 496-512, 1995  
A:Authors: Gnehm, C.L.; McDonald, L.A.; Small, K.V.; Fraser, C.M.; Smith, H.O.; Vente  
A:Title: Whole-genome random sequencing and assembly of Haemophilus influenzae Rd.  
A:Reference number: A64000; MUID:95350630  
A:Accession: G64086  
A>Status: nucleic acid sequence not shown; translation not shown  
A:Molecule type: DNA  
A:Residues: 1-364 <TIGR>  
A:Cross-references: GB:U32751; GB:L42023; NID:g3312200; PIDN:AAC22348.1; PID:g1573690  
A:Experimental source: strain Rd KW20  
C:Keywords: lipoprotein; phosphoric diester hydrolase; surface antigen

Query Match 24.4%; Score 553.5; DB 2: Length 364;  
Best Local Similarity 45.9%; Pred. No. 2.5e-35;  
Matches 141; Conservative 21; Mismatches 72; Indels 73; Gaps 11;

QY 4 SSHSSNNANTOMKSKDIIIAHRCASCYLPEHTLESKALAFQAQADYLEODLMTKDGRLV 63  
|||||  
DB 20 SSHSSNNANTOMKSKDIIIAHRCASCYLPEHTLESKALAFQAQADYLEODLMTKDGRLV 79  
|||||  
QY 64 VIHDFLDGLTDVAKKFPFHRKDGYYVIDFTLKEIQSLEMTENFETMGKK----- 115  
|||||  
DB 80 VIHDFLDGLTDVAKKFPFHRKDGYYVIDFTLKEIQSLEMTENFETMGKKQAQVYVNR 139  
|||||  
QY 116 -----W-----SKSSVVG-WPTVVRMRRAEPAADGVGAASRDLEK 150  
|||||  
DB 140 FPLWQSHFRHTFDEIEFIQGLEKSTGRKVGIYPEIKAPWFHONGKDIAETTLKVLLK 199  
|||||  
QY 151 HGAITSNTAATNAACAWLEAQAEEVEGFPVTPQVPLRPMTYKAAVDLSHFLKEKGGLEG 210  
|||||  
DB 200 YGYDKKTD-----MVLQTFDFNELKRIKTELLPQMGMDLK-LVQLIAYTDWK----- 246  
|||||  
QY 211 LIHSORRQDILDLIYHTOGYFPDQW-----NYTPGQVRYPLTFGWCYKLVVPE 260  
|||||  
DB 247 ---ETOEKDPKGYVW---NINYDMWFKPGMAEVVYKADGVGP-----GW-YMLVNKE 292  
|||||  
QY 261 ---PDKV 264  
|||  
DB 293 ESKPDNI 299

RESULT 28  
S59936  
glycerophosphodiester phosphodiesterase (EC 3.1.4.46) - Haemophilus influenzae (strai  
N:Alternate names: Igp-binding protein; protein D  
C:Species: Haemophilus influenzae  
A:Variety: strain NCTC 8468  
C:Date: 15-Feb-1996 #sequence\_revision 01-Mar-1996 #text\_change 08-Oct-1999  
C:Accession: S59936; S47338  
R:Song, X.M.; Forsgren, A.; Janson, H.  
Infect. Immun. 63, 696-699, 1995  
A:Title: The gene encoding protein D (hpd) is highly conserved among Haemophilus infl  
A:Reference number: S59931; MUID:95122210

A:Accession: S59936  
A:Molecule type: DNA  
A:Residues: 1-364 <SON>  
A:Cross-references: EMBL:235661; NID:9525223; PIDN:CAA84720.1; PID:9525224  
A:Experimental source: strain NCTC 8468  
C:Keywords: lipoprotein; phosphoric diester hydrolase; surface antigen

Query Match 24.2%; Score 547.5; DB 2; Length 364;  
Best Local Similarity 45.6%; Pred. No. 7.4e-35;  
Matches 140; Conservative 20; Mismatches 74; Indels 73; Gaps 11;  
  
Qy 4 SSSSSMANTOMKSKIIIAHAGSGLPEHTLESKALFAQAQADYLEODLMTKDGRLV 63  
Db 20 SSSSSMANTOMKSKIIIAHAGSGLPEHTLESKALFAQAQADYLEODLMTKDGRLV 79  
Qy 64 VIHDFLDGLTDAVAKKFFPHRRKDGYYVIDFTLKEIQSLEMTENFTMGKK----- 115  
Db 80 VIHDFLDGLTDAVAKKFFPHRRKDGYYVIDFTLKEIQSLEMTENFTMGKKQAQVYVNR 139  
Qy 116 ---W-----SKSVYG-WPTVRRMRRAPADGVGAASRLQEK 150  
Db 140 PFLMKSHFRIHTEDEIEFIOGLEKSTGKVGIYPEIKAPWFHONGKIDIAAETLKVLEK 199  
Qy 151 HGAITSSNTAATNACANLEAEVEEVPVTPQVPLRPMYKAAVDLSHFLKEKGGLEG 210  
Db 200 GYDKKTD-----MYLQTFDFNELAKRIKTELLPQMGMDLK-LVQLIATYDWR----- 246  
Qy 211 LIHSORRODILDLYHTTQGYFPDQ-----NYTPGPGVRYPLTFGWCYKLVPE 260  
Db 247 ---ETOEKDPKGYW----NYYNDMFKPGMAEVVYADGVGP-----GW-YMLVKNK 292  
Qy 261 ---PKV 264  
Db 293 ESKPDNI 299

RESULT 29

S33982

trans-activating transcription regulator - human immunodeficiency virus type 1  
C:Species: human immunodeficiency virus type 1, HIV-1  
C:Date: 06-Oct-1994 #sequence\_revision 23-Feb-1996 #text\_change 20-Sep-1999  
C:Accession: S33982; S26385; S19864  
R:Carlini, F.  
submitted to the EMBL Data Library, November 1991  
A:Reference number: S33979  
A:Accession: S33982  
A:Molecule type: DNA  
A:Residues: 1-86 <CAR>  
A:Cross-references: EMBL:211530; NID:960192; PIDN:CAA77625.1; PID:960196  
R:Siderovski, D.P.; Matsuyama, T.; Frigerio, E.; Chui, S.; Min, X.; Erfle, H.; Sumner-Sm  
Nucleic Acids Res. 20, 5311-5320, 1992  
A:Title: Random mutagenesis of the human immunodeficiency virus type-1 trans-activator c  
A:Reference number: S26385; MUID:93065196  
A:Accession: S26385  
A:Molecule type: nucleic acid  
A:Residues: 1-86 <SID>  
A:Cross-references: EMBL:X64650; NID:960144; PIDN:CAA5921.1; PID:960145  
C:Genetics:  
A:Gene: tat  
A:Introns: 72/2  
C:Superfamily: AIDS trans-activating transcription regulator  
C:Keywords: AIDS; immunodeficiency

Query Match 21.5%; Score 486; DB 2; Length 86;  
Best Local Similarity 100.0%; Pred. No. 7.4e-31;  
Matches 85; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 320 EPVDPRLPEWPKHPSQPKTACTNCCYKCCFHCQVCFITKALGISYGRKKRRRRPQG 379  
Db 2 EPVDPRLPEWPKHPSQPKTACTNCCYKCCFHCQVCFITKALGISYGRKKRRRRPQG 61

Qy 380 SOTHQVSLSKQPTSQSRGDPGTGPKE 404  
Db 62 SOTHQVSLSKQPTSQSRGDPGTGPKE 86

RESULT 30  
TNLJL2

trans-activating transcription regulator - human immunodeficiency virus type 1 (isol  
C:Species: human immunodeficiency virus type 1, HIV-1  
A:Note: host Homo sapiens (man)  
C:Date: 04-Dec-1986 #sequence\_revision 04-Dec-1986 #text\_change 02-Jul-1998  
C:Accession: A04017  
R:Arya, S.K.; Gallo, R.C.  
Proc. Natl. Acad. Sci. U.S.A. 83, 2209-2213, 1986  
A:Title: Three novel genes of human T-lymphotropic virus type III: immune reactivity  
A:Reference number: A94093; MUID:86177573  
A:Accession: A04017  
A:Molecule type: DNA  
A:Residues: 1-95 <ARY>  
C:Genetics:  
A:Gene: tat  
C:Superfamily: AIDS trans-activating transcription regulator  
C:Keywords: AIDS; immunodeficiency; transcription regulation

Query Match 21.1%; Score 478; DB 1; Length 95;  
Best Local Similarity 98.8%; Pred. No. 3.5e-30;  
Matches 84; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy 320 EPVDPRLPEWPKHPSQPKTACTNCCYKCCFHCQVCFITKALGISYGRKKRRRRPQG 379  
Db 11 EPVDPRLPEWPKHPSQPKTACTNCCYKCCFHCQVCFITKALGISYGRKKRRRRPQG 70

Qy 380 SOTHQVSLSKQPTSQSRGDPGTGPKE 404  
Db 71 SOTHQVSLSKQPTSQSRGDPGTGPKE 95

Search completed: August 26, 2002, 08:13:23  
Job time: 292 sec

GenCore version 4.5  
Copyright (c) 1993 - 2000 Compugen Ltd.

OM protein - protein search, using sw model

Run on: August 26, 2002, 08:10:40 ; Search time 38.46 Seconds  
(without-alignments)  
262.293 Million cell updates/sec

Title: US-09-509-239-21  
Perfect score: 2264  
Sequence: 1 MDPSSSSNWMANTQMSDKI.....QSRGDTGPKETSGHHHHH 413

Scoring table:  
BLOSUM62  
Gapop 10.0 , Gapext 0.5

Searched: 231628 seqs, 24425594 residues

Total number of hits satisfying chosen parameters: 231628

Minimum DB seq length: 0  
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

Database : Issued Patents AA.\*  
1: /cgn2\_6/ptodata/2/1aa/5A-COMB.pap.\*  
2: /cgn2\_6/ptodata/2/1aa/5B-COMB.pap.\*  
3: /cgn2\_6/ptodata/2/1aa/6A-COMB.pap.\*  
4: /cgn2\_6/ptodata/2/1aa/6B-COMB.pap.\*  
5: /cgn2\_6/ptodata/2/1aa/PCTUS-COMB.pap.\*  
6: /cgn2\_6/ptodata/2/1aa/backfiles1.pap.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	1138	50.3	239	4	US-08-679-493A-76
2	1115	49.2	206	3	US-08-388-353-644
3	1115	49.2	206	3	US-08-488-551B-644
4	971	42.9	206	4	US-08-679-493A-75
5	638	28.2	123	4	US-09-124-900-10
6	603.5	26.7	278	4	US-09-485-885-21
7	597	26.4	383	4	US-09-485-885-23
8	594	26.2	220	4	US-09-485-885-8
9	593	26.2	220	4	US-09-485-885-1
10	590	26.1	227	4	US-09-485-885-16
11	590	26.1	371	4	US-09-485-885-6
12	587	25.9	227	4	US-09-485-885-19
13	582	25.7	273	4	US-09-485-885-4
14	498.5	22.0	151	4	US-08-679-493A-74
15	486	21.5	86	1	US-08-450-237-1
16	486	21.5	86	1	US-08-450-246-1
17	486	21.5	86	1	US-08-450-098-1
18	486	21.5	86	1	US-08-451-233-1
19	486	21.5	86	1	US-08-450-236-1
20	486	21.5	86	3	US-07-808-452-1
21	486	21.5	86	3	US-09-030-613-19
22	486	21.5	86	4	US-09-124-900-6
23	486	21.5	86	4	US-09-451-905-19
24	486	21.5	86	4	US-08-235-403-1
25	486	21.5	86	5	PCT-US92-10770-1
26	486	21.5	86	5	PCT-US95-06077-2
27	478.5	21.1	266	3	US-08-815-809-5

28	478.5	21.1	280	2	US-08-816-155B-43	Sequence 43, Appl
29	478.5	21.1	280	4	US-09-079-587-43	Sequence 43, Appl
30	470	20.8	86	2	US-08-505-210-1	Sequence 1, Appl
31	470	20.8	86	4	US-09-099-333-1	Sequence 1, Appl
32	466	20.6	83	2	US-08-417-210A-96	Sequence 96, Appl
33	455	20.1	82	1	US-08-053-079A-15	Sequence 15, Appl
34	439	19.4	253	2	US-08-659-251-4	Sequence 4, Appl
35	439	19.4	253	4	US-09-256-490-4	Sequence 4, Appl
36	439	19.4	253	5	PCT-US96-11445-4	Sequence 4, Appl
37	409	18.1	72	3	US-09-030-613-17	Sequence 17, Appl
38	409	18.1	72	4	US-09-451-905-17	Sequence 17, Appl
39	385	17.0	72	2	US-08-893-853-1	Sequence 1, Appl
40	385	17.0	72	4	US-09-113-921-1	Sequence 1, Appl
41	372	16.4	122	6	5304466-3	Patent No. 5304466
42	365.5	16.1	312	1	US-08-094-128A-27	Sequence 27, Appl
43	365.5	16.1	312	1	US-08-455-674-27	Sequence 27, Appl
44	365.5	16.1	312	1	US-08-455-992-27	Sequence 27, Appl
45	365.5	16.1	312	1	US-08-455-972-27	Sequence 27, Appl

ALIGNMENTS

RESULT 1  
US-08-679-493A-76  
; Sequence 76, Application US/08679493A  
; Patent No. 6303295  
; GENERAL INFORMATION:  
; APPLICANT: Taylor, Ethan W.  
; TITLE OF INVENTION: SELENOPROTEINS, CODING SEQUENCES AND METHODS  
; FILE REFERENCE: 55-95  
; CURRENT APPLICATION NUMBER: US/08/679,493A  
; CURRENT FILING DATE: 1996-07-12  
; PRIOR APPLICATION NUMBER: 60/001203  
; PRIOR FILING DATE: 1995-07-14  
; PRIOR APPLICATION NUMBER: 60/003,112  
; PRIOR FILING DATE: 1995-09-01  
; NUMBER OF SEQ ID NOS: 216  
; SOFTWARE: PatentIn Ver. 2.0  
; SEQ ID NO 76  
; LENGTH: 239  
; TYPE: PRT  
; ORGANISM: Human immunodeficiency virus type 1  
; FEATURE:  
; NAME/KEY: VARIANT  
; LOCATION: (1)..(239)  
; OTHER INFORMATION: X os selenocysteine.  
US-08-679-493A-76

Query Match	50.3%	Score 1138;	DB 4;	Length 239;
Best Local Similarity	88.7%	Pred. No. 5e-107;		
Matches 211;	Conservative	2;	Mismatches 25;	Indels 0;
Gaps	0;			
QY	112	MGGKSKSVVWCPVPRMRRAEPAADGVGAASRDLEKHGAITSNTAATNAACAWLEA	171	
DB	1	MGGKSKSVVWCPVPRMRRAEPAADGVGAASRDLEKHGAITSNTAATNAACAWLEA	60	
QY	172	QEEVEGFPVTPQVPLRPMYKAAVDLSHFLKEGKGLIHSORRQDILDWIYHTQGY	231	
DB	61	QEEVEGFPVTPQVPLRPMYKAAVDLSHFLKEGKGLIHSORRQDILDWIYHTQGY	120	
QY	232	FPDWNYPGPGVRYPLTFGWCYKLVVPDPKVEEANKGENTSLHPVSLHGMDDPEREV	291	
DB	121	FPDWNYPGPGVRYPLTFGWCYKLVVPDPKVEEANKGENTSLHPVSLHGMDDPEREV	180	
QY	292	LEWRDSDRLAFHHVARELHPEYFKNCTSEPVDPRLPEWKHPGSPKTACTNCKYKCC	349	
DB	181	LEWRDSDRLAFHHVARELHPEYFKNCTSEPVDPRLPEWKHPGSPKTACTNCKYKCC	238	

RESULT 2  
US-08-388-353-644

Sequence 644, Application US/08388353  
Patent No. 6010895  
GENERAL INFORMATION:  
APPLICANT: Deacon, Nicholas J.  
APPLICANT: Learnmont, Jennifer C.  
APPLICANT: McPhee, Dale A.  
APPLICANT: Crowe, Suzanne  
APPLICANT: Cooper, David  
TITLE OF INVENTION: NON-PATHOGENIC STRAINS OF HIV-1  
NUMBER OF SEQUENCES: 800  
CORRESPONDENCE ADDRESSES:  
ADDRESSEE: Scully, Scott, Murphy & Presser  
STREET: 400 Garden City Plaza  
CITY: Garden City  
STATE: New York  
COUNTRY: United States  
ZIP: 11530  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: Patentin Release #1.0, Version #1.25  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/388,353  
FILING DATE: 14-FEB-1995  
CLASSIFICATION: 424  
ATTORNEY/AGENT INFORMATION:  
NAME: Digiglio, Frank S.  
REGISTRATION NUMBER: 31,346  
REFERENCE/DOCKET NUMBER: 9606  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: (516) 742-4343  
TELEFAX: (516) 742-4366  
TELEX: 230 901 SANS UR  
INFORMATION FOR SEQ ID NO: 644:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 206 amino acids  
TYPE: amino acid  
STRANDEDNESS: single  
TOPOLOGY: linear  
MOLECULE TYPE: protein  
US-08-388-353-644

Query Match 49.28; Score 1115; DB 3; Length 206;  
Best Local Similarity 98.18; Pred. No. 8.5e-105;  
Matches 202; Conservative 1; Mismatches 3; Indels 0; Gaps 0;  
Qy 112 MGGKWSKSSVVGWPTVRRMRRAEPAADGGAASRDLEKKGGLIHSQRQDILDIWIYHTQY 171  
Db 1 MGGKWSKSSVVGWPAVRMRRAEPAADGGAASRDLEKKGGLIHSQRQDILDIWIYHTQY 60  
Qy 172 QEEVEGFPVTPQVPLRPMTYKAAVDLSHFLKEKGGLEGLIHSQRQDILDIWIYHTQY 231  
Db 61 QEEVEGFPVTPQVPLRPMTYKAAVDLSHFLKEKGGLEGLIHSQRQDILDIWIYHTQY 120  
Qy 232 FPDQNYTPGCVRYPLTFGCYKLVPEPKVEANKGENTSLHHPVSLHGMDDPEREV 291  
Db 121 FPDQNYTPGCVRYPLTFGCYKLVPEPKVEANKGENTSLHHPVSLHGMDDPEREV 180  
Qy 292 LEWRFDLSRLAFHHVARELHPEYFKNC 317  
Db 181 LEWRFDLSRLAFHHVARELHPEYFKNC 206

RESULT 3  
US-08-488-551B-644  
Sequence 644, Application US/08488551B  
Patent No. 6015661  
GENERAL INFORMATION:  
APPLICANT: Nicholas J. Deacon  
APPLICANT: Dale A. McPhee  
APPLICANT: David Cooper

TITLE OF INVENTION: NON-PATHOGENIC STRAINS OF HIV-1  
NUMBER OF SEQUENCES: 841  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: SCULLY, SCOTT, MURPHY & PRESSER  
STREET: 400 GARDEN CITY PLAZA  
CITY: GARDEN CITY  
STATE: NEW YORK  
COUNTRY: U.S.A.  
ZIP: 11530-0299  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: Patentin Release #1.0, Version #1.25  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/488,551B  
FILING DATE: 07-JUN-1995  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: PM3864 (AU)  
FILING DATE: 14-FEB-1994  
APPLICATION NUMBER: PM4002 (AU)  
FILING DATE: 21-FEB-1994  
APPLICATION NUMBER: PM0284 (AU)  
FILING DATE: 23-DEC-1994  
APPLICATION NUMBER: US 08/388,353  
FILING DATE: 14-FEB-1995  
APPLICATION NUMBER: PM3021/95  
FILING DATE: 17-MAY-1995  
ATTORNEY/AGENT INFORMATION:  
NAME: FRANK S. DIGIGLIO  
REFERENCE/DOCKET NUMBER: 96062  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: (516) 742-4343  
TELEFAX: (516) 742-4366  
INFORMATION FOR SEQ ID NO: 644:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 206 amino acids  
TYPE: amino acid  
STRANDEDNESS: single  
TOPOLOGY: linear  
MOLECULE TYPE: protein  
US-08-488-551B-644

Query Match 49.28; Score 1115; DB 3; Length 206;  
Best Local Similarity 98.18; Pred. No. 8.5e-105;  
Matches 202; Conservative 1; Mismatches 3; Indels 0; Gaps 0;  
Qy 112 MGGKWSKSSVVGWPTVRRMRRAEPAADGGAASRDLEKKGGLIHSQRQDILDIWIYHTQY 171  
Db 1 MGGKWSKSSVVGWPAVRMRRAEPAADGGAASRDLEKKGGLIHSQRQDILDIWIYHTQY 60  
Qy 172 QEEVEGFPVTPQVPLRPMTYKAAVDLSHFLKEKGGLEGLIHSQRQDILDIWIYHTQY 231  
Db 61 QEEVEGFPVTPQVPLRPMTYKAAVDLSHFLKEKGGLEGLIHSQRQDILDIWIYHTQY 120  
Qy 232 FPDQNYTPGCVRYPLTFGCYKLVPEPKVEANKGENTSLHHPVSLHGMDDPEREV 291  
Db 121 FPDQNYTPGCVRYPLTFGCYKLVPEPKVEANKGENTSLHHPVSLHGMDDPEREV 180  
Qy 292 LEWRFDLSRLAFHHVARELHPEYFKNC 317  
Db 181 LEWRFDLSRLAFHHVARELHPEYFKNC 206

RESULT 4  
US-08-679-493A-75  
Sequence 75, Application US/08679493A  
Patent No. 6303295  
GENERAL INFORMATION:  
APPLICANT: Taylor, Ethan W.  
TITLE OF INVENTION: SELENOPROTEINS, CODING SEQUENCES AND METHODS  
FILE REFERENCE: 55-95

;; CURRENT APPLICATION NUMBER: US/08/679,493A  
;; CURRENT FILING DATE: 1996-07-12  
;; PRIOR APPLICATION NUMBER: 60/001203  
;; PRIOR FILING DATE: 1995-07-14  
;; PRIOR APPLICATION NUMBER: 60/003,112  
;; PRIOR FILING DATE: 1995-09-01  
;; NUMBER OF SEQ ID NOS: 216  
;; SOFTWARE: PatentIn Ver. 2.0  
;; SEQ ID NO 75  
;; LENGTH: 206  
;; TYPE: PRT  
;; ORGANISM: Human immunodeficiency virus type 1  
;; NAME/KEY: VARIANT  
;; FEATURE:  
;; LOCATION: (1)..(206)  
;; OTHER INFORMATION: X is selenocysteine.  
US-08-679-493A-75

Query Match 42.9%; Score 971; DB 4; Length 206;  
Best Local Similarity 89.0%; Pred. No. 3e-90; Indels 6; Gaps 3;  
Matches 186; Conservative 2; Mismatches 15;  
Qy 112 MGGKWSKSSVVGWPTVRMRRAEPAADGGAASRDLEKHGAITSNTAATNAACAWLEA 171  
Db 1 MGGKWSKSSVVGWPTVRMRRAEPAADGGAASRDLEKHGAITSNTAATNAACAWLEA 60  
Qy 172 QEEVEGFPVTPQVPLRPMYKAAVDLSHFLKEKGG---LEGLIHSQRRQDILDWIYHT 228  
Db 61 QEEVEGFPVTPQVPLRPMYKAAVDLSHFLKRGKGTGRANSLPTKTPYXSVDL--PHT 118  
Qy 229 QCYFDMQNYTPGQVRYPLTFGWCYKLVPEPDKVEEANKGENTSLHPVSLHGMDDPE 288  
Db 119 R-LLPDMQNYTPGQVRYPLTFGWCYKLVPEPDKVEEANKGENTSLHPVSLHGMDDPE 177  
Qy 289 REVLEWRFSRLAFHFAHVAELHPEYFKNC 317  
Db 178 REVLEWRFSRLAFHFAHVAELHPEYFKNC 206

RESULT 5  
US-09-124-900-10  
;; Sequence 10, Application US/09124900  
;; Patent No. 6268484  
;; GENERAL INFORMATION:  
;; APPLICANT: KATINGER, Hermann  
;; APPLICANT: BUCHACHER, Andrea  
;; APPLICANT: ERNST, Wolfgang  
;; APPLICANT: BALLAUN, Claudia  
;; APPLICANT: PURTSCHER, Martin  
;; APPLICANT: TRKOLA, Alexandra  
;; APPLICANT: FREDL, Renate  
;; APPLICANT: SCHMATZ, Christine  
;; APPLICANT: KLINA, Annelies  
;; APPLICANT: STEINDL, Franz  
;; APPLICANT: MUSTER, Thomas  
;; TITLE OF INVENTION: HIV-vaccines  
;; FILE REFERENCE: 1939-112P  
;; CURRENT APPLICATION NUMBER: US/09/124,900  
;; CURRENT FILING DATE: 1998-07-30  
;; PRIOR APPLICATION NUMBER: PCT/EP95/01481  
;; PRIOR FILING DATE: 1995-04-19  
;; NUMBER OF SEQ ID NOS: 11  
;; SOFTWARE: PatentIn version 3.0  
;; SEQ ID NO 10  
;; LENGTH: 123  
;; TYPE: PRT  
;; ORGANISM: Human immunodeficiency virus type 1  
US-09-124-900-10

Query Match 28.2%; Score 638; DB 4; Length 123;  
Best Local Similarity 97.6%; Pred. No. 6.5e-57;

Matches 120; Conservative 0; Mismatches 3; Indels 0; Gaps 0;  
Qy 112 MGGKWSKSSVVGWPTVRMRRAEPAADGGAASRDLEKHGAITSNTAATNAACAWLEA 171  
Db 1 MGGKWSKSSVVGWPTVRMRRAEPAADGGAASRDLEKHGAITSNTAATNAACAWLEA 60  
Qy 172 QEEVEGFPVTPQVPLRPMYKAAVDLSHFLKEKGGLEGLIHSQRRQDILDWIYHTGY 231  
Db 61 QEEVEGFPVTPQVPLRPMYKAAVDLSHFLKEKGGLEGLIHSQRRQDILDWIYHTGY 120  
Qy 232 FPD 234  
Db 121 FPD 123  
RESULT 6  
US-09-485-885-21  
;; Sequence 21, Application US/09485885  
;; Patent No. 6342224  
;; GENERAL INFORMATION:  
;; APPLICANT: Bruck, Claudine  
;; APPLICANT: Cabezon Silva, Teresa  
;; APPLICANT: Delisse, Anne-Marie Eva Fernande  
;; APPLICANT: Gerard, Catherine Marie Ghislaine  
;; APPLICANT: Lombardo-Bencheikh, Angela  
;; TITLE OF INVENTION: Vaccine  
;; FILE REFERENCE: B45107  
;; CURRENT APPLICATION NUMBER: US/09/485,885  
;; CURRENT FILING DATE: 2000-02-18  
;; PRIOR APPLICATION NUMBER: PCT/EP98/05285  
;; PRIOR FILING DATE: 1998-08-17  
;; PRIOR APPLICATION NUMBER: GB 9717953.5  
;; PRIOR FILING DATE: 1997-08-22  
;; NUMBER OF SEQ ID NOS: 23  
;; SOFTWARE: FastSeq for Windows Version 3.0  
;; SEQ ID NO 21  
;; LENGTH: 278  
;; TYPE: PRT  
;; ORGANISM: Homo sapien  
US-09-485-885-21

Query Match 26.7%; Score 603.5; DB 4; Length 278;  
Best Local Similarity 36.4%; Pred. No. 6.7e-53;  
Matches 158; Conservative 26; Mismatches 73; Indels 177; Gaps 12;  
Qy 1 MDPSSSHSNMANTQMSDKIIITAHRGASGYLPEHTLESKALAFQAQADYLEODLMTKDG 60  
Db 1 MDPSSSHSNMANTQMSDKIIITAHRGASGYLPEHTLESKALAFQAQADYLEODLMTKDG 60  
Qy 61 RLWVIHDFLGLTQVAKKFPHRHRKDGRIYVIDFTLKEIQSLEMTENPFTMGKWSKSS 120  
Db 61 RLWVIHDFLGLTQVAKKFPHRHRKDGRIYVIDFTLKEIQSLEMTENPFTMA----- 113  
Qy 121 VVGWPTVRMRRAEPAADGGAASRDLEKHGAITSNTAATN--AACAW----LEAOEE 174  
Db 114 -----RFEDP-----TRPYKLPDLCELNTSLQDIEITCYICKTVLLETEV 155  
Qy 175 EEVGFPVTPQVPLRPMYKAAVDLSHFLKEKGGLEGLIHSQRRQDILDWIYHTGYFPD 234  
Db 156 REFAFK-----DLFVVY----- 167  
Qy 235 WONYTPGQVRYPLTFGWCYKLV-----PVEPDKVEE-ANKGENTSLHPV 279  
Db 168 -----RDSIPHAACHKCIDFYSRIRELRHYSDSYGDTLEKLTNTGLYNLLIRCL 217  
Qy 280 SLHGMDPEREVLEWRFSRLAFHFAHVAELHPEYFKNCTSEPVDPRLPWPKHGPGOPKTA 339  
Db 218 RCQKPLNPAEKLR--HLNEKRRFHNIAGHYRCQ----- 248  
Qy 340 CTNICYCKKCCFHCQVCFITKALGISYGRKKRRQRRPPQGSQTHVSLKSQPTQSRRGP 399  
Db 249 -----CHSCC-----NRQERLQRRRETQ----- 268

QY 400 TGPKESTGHHHHH 413  
Db 269 -----VTSGHHHHH 278

RESULT 7  
US-09-485-885-23  
; Sequence 23, Application US/09485885  
; Patent No. 6342224  
; GENERAL INFORMATION:  
; APPLICANT: Bruck, Claudine  
; APPLICANT: Cabezon Silva, Teresa  
; APPLICANT: Delisse, Anne-Marie Eva Fernande  
; APPLICANT: Gerard, Catherine Marie Ghislaine  
; APPLICANT: Lombardo-Bencheikh, Angela  
; TITLE OF INVENTION: Vaccine  
; FILE REFERENCE: B45107  
; CURRENT APPLICATION NUMBER: US/09/485.885  
; CURRENT FILING DATE: 2000-02-18  
; PRIOR APPLICATION NUMBER: PCT/EP98/05285  
; PRIOR FILING DATE: 1998-08-17  
; PRIOR APPLICATION NUMBER: GB 9717953.5  
; PRIOR FILING DATE: 1997-08-22  
; NUMBER OF SEQ ID NOS: 23  
; SOFTWARE: FastSeq for Windows Version 3.0  
; SEQ ID NO 23  
; LENGTH: 383  
; TYPE: PRT  
; ORGANISM: Homo sapien  
US-09-485-885-23

Query Match 26.4%; Score 597; DB 4; Length 383;  
Best Local Similarity 34.9%; Pred. No. 4.8e-52;  
Matches 170; Conservative 21; Mismatches 118; Indels 178; Gaps 11;  
QY 1 MDPSSSHSNMANTOMKSKIIIAHARGASGYLPEHTLESKALAFQAQADYLEODLMTKDG 60  
Db 1 MDPSSSHSNMANTOMKSKIIIAHARGASGYLPEHTLESKALAFQAQADYLEODLMTKDG 60  
QY 61 RLWVIHDFLDGLTDVAKKFPHRHRKDGYYVIDFTLKEIOSLEMTENFETMGGKSS 114  
Db 61 RLWVIHDFLDGLTDVAKKFPHRHRKDGYYVIDFTLKEIOSLEMTENFETMGGKSS 120  
QY 115 -----KWSKSSVGVWPTVRRMRRAEPAADGVGAASRDLEKHAITS-- 156  
Db 121 RPKLPDLCTELNTSLQDIEITCVYCKTVLETFEFAFKDLFVYVYRDSIPHAACHKCI 180  
QY 157 -----SNTAATNAACAWLEAQEEVEGFPVTPQVPLRPMY 192  
Db 181 DFYSRIRELHYDSVYGDTEKLTNTGLYNLLIRLCRCQK-----PLNPAEKLRLHNE 234  
QY 193 KAADLSHFLKEKGGLGLHS---QRRQDILDWIYHTQGYFPDQWQNYTEPGVRYPLT 249  
Db 235 K-----RRFNTIAGHYRGQCHCCNRARQ----- 258  
QY 250 FGWCYKLPVPEPKVEEANKGENTSLLH-----PVSL---HGMDDP 287  
Db 259 -----ERLQRRRETQVMHGPKATLDQIVLHLEPQNEIPVDLLCHEQLSDS 303  
QY 288 EREVLEWRDRLAPHVHARELHPE-YFNKCTSEPVDRLEPWKHPGSPKACTNCYCK 346  
Db 304 EGENDE--IDGVNHOHLPARRAPORHTMLCMCCKEARIELVWESSADDLRAFOQLFN 361  
QY 347 KCCFCQCVFITKALGISYGRKKRRRPPQSGQTHOVSLSKQPTSOSRGDPTGPKETS 406  
Db 362 TLSFVCPWC-----ASQOTS 376  
QY 407 GHHHHHH 413  
Db 377 GHHHHHH 383

RESULT 8  
US-09-485-885-8  
; Sequence 8, Application US/09485885  
; Patent No. 6342224  
; GENERAL INFORMATION:  
; APPLICANT: Bruck, Claudine  
; APPLICANT: Cabezon Silva, Teresa  
; APPLICANT: Delisse, Anne-Marie Eva Fernande  
; APPLICANT: Gerard, Catherine Marie Ghislaine  
; APPLICANT: Lombardo-Bencheikh, Angela  
; TITLE OF INVENTION: Vaccine  
; FILE REFERENCE: B45107  
; CURRENT APPLICATION NUMBER: US/09/485.885  
; CURRENT FILING DATE: 2000-02-18  
; PRIOR APPLICATION NUMBER: PCT/EP98/05285  
; PRIOR FILING DATE: 1998-08-17  
; PRIOR APPLICATION NUMBER: GB 9717953.5  
; PRIOR FILING DATE: 1997-08-22  
; NUMBER OF SEQ ID NOS: 23  
; SOFTWARE: FastSeq for Windows Version 3.0  
; SEQ ID NO 8  
; LENGTH: 220  
; TYPE: PRT  
; ORGANISM: Homo sapien  
US-09-485-885-8

Query Match 26.2%; Score 594; DB 4; Length 220;  
Best Local Similarity 67.0%; Pred. No. 4.3e-52;  
Matches 126; Conservative 9; Mismatches 27; Indels 26; Gaps 3;  
QY 1 MDPSSSHSNMANTOMKSKIIIAHARGASGYLPEHTLESKALAFQAQADYLEODLMTKDG 60  
Db 1 MDPSSSHSNMANTOMKSKIIIAHARGASGYLPEHTLESKALAFQAQADYLEODLMTKDG 60  
QY 61 RLWVIHDFLDGLTDVAKKFPHRHRKDGYYVIDFTLKEIOSLEMTENFETMGGKSS 120  
Db 61 RLWVIHDFLDGLTDVAKKFPHRHRKDGYYVIDFTLKEIOSLEMTENFETMGGKSS 118  
QY 121 VGVNPTVRRMRRAEPAADGVGAASRDLEKHAITSNTAATNAACAWLEAQEEVEGFP 180  
Db 119 -----PTLHEYMLDLP-----ETDLYGQQLNDS-----SEEEDEIDGP 154  
QY 181 VTPQVPLR 188  
Db 155 AGQAEPR 162  
RESULT 9  
US-09-485-885-1  
; Sequence 1, Application US/09485885  
; Patent No. 6342224  
; GENERAL INFORMATION:  
; APPLICANT: Bruck, Claudine  
; APPLICANT: Cabezon Silva, Teresa  
; APPLICANT: Delisse, Anne-Marie Eva Fernande  
; APPLICANT: Gerard, Catherine Marie Ghislaine  
; APPLICANT: Lombardo-Bencheikh, Angela  
; TITLE OF INVENTION: Vaccine  
; FILE REFERENCE: B45107  
; CURRENT APPLICATION NUMBER: US/09/485.885  
; CURRENT FILING DATE: 2000-02-18  
; PRIOR APPLICATION NUMBER: PCT/EP98/05285  
; PRIOR FILING DATE: 1998-08-17  
; PRIOR APPLICATION NUMBER: GB 9717953.5  
; PRIOR FILING DATE: 1997-08-22  
; NUMBER OF SEQ ID NOS: 23  
; SOFTWARE: FastSeq for Windows Version 3.0  
; SEQ ID NO 1  
; LENGTH: 220  
; TYPE: PRT  
; ORGANISM: Homo sapien



US-09-485-885-1

Query Match 26.2%; Score 593; DB 4; Length 220;  
Best Local Similarity 67.0%; Pred. No. 5.4e-52;  
Matches 126; Conservative 9; Mismatches 27; Indels 26; Gaps 3;

QY 1 MDPSSSHSNANTQMSDKIIIAHRCASGYLPEHTLESKALAFQAQADYLEODLAWTKDG 60  
DB 1 MDPSSSHSNANTQMSDKIIIAHRCASGYLPEHTLESKALAFQAQADYLEODLAWTKDG 60

QY 61 RLWVIHDFDLGLTDVAKKFPFHRKRDGRYYVIDFTLKEIQSLEMTENFETMGKWSKSS 120  
DB 61 RLWVIHDFDLGLTDVAKKFPFHRKRDGRYYVIDFTLKEIQSLEMTENFETMGKWSKSS 118

QY 121 VVGWPTVRMRRAEPAADGVGAASRDLEKHGAISSNTAATNAACAWLEAQEEVEGPP 180  
DB 119 ----PTLHEYMLDLP-----ETDLYCYEQLNDS-----SEEDIEDGP 154

QY 181 VTPOVPLR 188  
DB 155 AGQAEPR 162

RESULT 10

US-09-485-885-16  
; Sequence 16, Application US/09485885  
; Patent No. 6342224  
; GENERAL INFORMATION:  
; APPLICANT: Bruck, Claudine  
; APPLICANT: Cabezon Silva, Teresa  
; APPLICANT: Delisse, Anne-Marie Eva Fernande  
; APPLICANT: Gerard, Catherine Marie Ghislaine  
; APPLICANT: Lombardo-Bencheikh, Angela  
; TITLE OF INVENTION: Vaccine  
; FILE REFERENCE: B45107  
; CURRENT FILING DATE: 2000-02-18  
; PRIOR APPLICATION NUMBER: US/09/485,885  
; PRIOR FILING DATE: 1998-08-17  
; PRIOR APPLICATION NUMBER: PCT/EP98/05285  
; PRIOR FILING DATE: 1998-08-17  
; PRIOR APPLICATION NUMBER: GB 9717953.5  
; NUMBER OF SEQ ID NOS: 23  
; SOFTWARE: FastSeq for Windows Version 3.0  
; SEQ ID NO 16  
; LENGTH: 227  
; TYPE: PRT  
; ORGANISM: Homo sapien  
US-09-485-885-16

Query Match 26.1%; Score 590; DB 4; Length 227;  
Best Local Similarity 37.0%; Pred. No. 1.1e-51;  
Matches 153; Conservative 14; Mismatches 59; Indels 188; Gaps 10;

QY 1 MDPSSSHSNANTQMSDKIIIAHRCASGYLPEHTLESKALAFQAQADYLEODLAWTKDG 60  
DB 1 MDPSSSHSNANTQMSDKIIIAHRCASGYLPEHTLESKALAFQAQADYLEODLAWTKDG 60

QY 61 RLWVIHDFDLGLTDVAKKFPFHRKRDGRYYVIDFTLKEIQSLEMTENFETMGKWSKSS 120  
DB 61 RLWVIHDFDLGLTDVAKKFPFHRKRDGRYYVIDFTLKEIQSLEMTENFETM----- 112

QY 121 VVGWPTVRMRRAEPAADGVGAASRDLEKHGAISSNTAATNAACAWLEAQEEVEGPP 180  
DB 113 -----AMHGPKATLQDIVLH-----LEPQNE----- 133

QY 181 VTPOVPLRMTYKAAVDLSHFLKEGGLTHSQRQDILDLWIYHTGYPDWQNTYPT 240  
DB 134 ----IP-----VDL-----LCHQELSDS----- 147

QY 241 GPCVRYPLTFGWCYKLVVPVEPKVEEANKGENTSLHHPVSLHGMDDPERVLEWRFDSRL 300

Db 148 -----EEENDEIDEVN-----HQ----- 160

QY 301 AFHHVARELHPE-YFKNCTSEPVDPRLPEWKHPGSPKTACTNCTCYCKKCCFHCQVCFITK 359  
DB 161 --HLPARRAEPQRHTMLCWCCKCEARIELVWSSADDLRAFOOLFNTLSFVCPWC----- 214

QY 360 ALGISYGRKKRQRPPGOSGTHOVSLSKPTQSQRGDPGPKTSGHHHHH 413  
DB 215 -----ASQQTSGHHHHH 227

RESULT 11  
US-09-485-885-6  
; Sequence 6, Application US/09485885  
; Patent No. 6342224  
; GENERAL INFORMATION:  
; APPLICANT: Bruck, Claudine  
; APPLICANT: Cabezon Silva, Teresa  
; APPLICANT: Delisse, Anne-Marie Eva Fernande  
; APPLICANT: Gerard, Catherine Marie Ghislaine  
; APPLICANT: Lombardo-Bencheikh, Angela  
; TITLE OF INVENTION: Vaccine  
; FILE REFERENCE: B45107  
; CURRENT APPLICATION NUMBER: US/09/485,885  
; CURRENT FILING DATE: 2000-02-18  
; PRIOR APPLICATION NUMBER: PCT/EP98/05285  
; PRIOR FILING DATE: 1998-08-17  
; PRIOR APPLICATION NUMBER: GB 9717953.5  
; PRIOR FILING DATE: 1997-08-22  
; NUMBER OF SEQ ID NOS: 23  
; SOFTWARE: FastSeq for Windows Version 3.0  
; SEQ ID NO 6  
; LENGTH: 371  
; TYPE: PRT  
; ORGANISM: Homo sapien  
US-09-485-885-6

Query Match 26.1%; Score 590; DB 4; Length 371;  
Best Local Similarity 34.9%; Pred. No. 2.4e-51;  
Matches 167; Conservative 27; Mismatches 112; Indels 172; Gaps 14;

QY 1 MDPSSSHSNANTQMSDKIIIAHRCASGYLPEHTLESKALAFQAQADYLEODLAWTKDG 60  
DB 1 MDPSSSHSNANTQMSDKIIIAHRCASGYLPEHTLESKALAFQAQADYLEODLAWTKDG 60

QY 61 RLWVIHDFDLGLTDVAKKFPFHRKRDGRYYVIDFTLKEIQSLEMTENFETMGKWSKSS 120  
DB 61 RLWVIHDFDLGLTDVAKKFPFHRKRDGRYYVIDFTLKEIQSLEMTENFETM----- 112

QY 121 VVGWPTVRMRRAEPAADGVGAASRDLEKHGAISSNTAATNAACAWLEAQEEVEGPP 180  
DB 113 -AMFQDPOERPKLPOLCTELQTTIHI-----ILECVYCKQ----- 149

QY 181 VTPOVPLRMTYKAAVDLSHFLKEGGLTHSQRQDILDLWIYHTG----- 230

DB 150 -----LLREVYDFA-----FRDLCTVYRDGNPYAVCDKCL 180

QY 231 -YFPDQWNTYPTGPGVRYPLTFGWCYKLVVPVEPKVEEANKGENTSLHHPVSLHGMDDPER 289  
DB 181 KYSKISEYR-----HYCSLYGTTLE--QYANKPLCDLLIRCINCKPLCPDE 227

QY 290 EYLEWRFDSRLAFHHVARELHPEYFNKCTSEPVDPRL-----PMKHP--GSPKTKAC 340  
DB 228 K--QRHLDKKQRFHNIIRGRTGRCMCCSSRTRRETQLMHGDTPTLHEYMLDLPETD 285

QY 341 TNCY-----CKKCFHCQVC----- 355

DB 286 LYCYEQALNDSSEEDIEDGPAGQAEPRAHYNIIVTFCKCDSTLRLCVQSTHVDITRLD 345

QY 356 FITKALGISYGRKKRQRPPGOSGTHOVSLSKPTQSQRGDPGPKTSGHHHHH 413  
DB 346 LLMGTLGIVC-----PICSQ-----KPTSGHHHHH 371

## RESULT 12

US-09-485-885-19  
; Sequence 19, Application US/09485885  
; Patent No. 6342224  
; GENERAL INFORMATION:  
; APPLICANT: Bruck, Claudine  
; APPLICANT: Cabezón Silva, Teresa  
; APPLICANT: Delisse, Anne-Marie Eva Fernande  
; APPLICANT: Gerard, Catherine Marie Ghislaine  
; APPLICANT: Lombardo-Bencheikh, Angela  
; TITLE OF INVENTION: Vaccine  
; FILE REFERENCE: B45107  
; CURRENT APPLICATION NUMBER: US/09/485,885  
; CURRENT FILING DATE: 2000-02-18  
; PRIOR APPLICATION NUMBER: PCT/EP98/05285  
; PRIOR FILING DATE: 1998-08-17  
; PRIOR APPLICATION NUMBER: GB 9717953.5  
; PRIOR FILING DATE: 1997-08-22  
; NUMBER OF SEQ ID NOS: 23  
; SOFTWARE: FastSeq for Windows Version 3.0  
; SEQ ID NO 19  
; LENGTH: 227  
; TYPE: PRT  
; ORGANISM: Homo sapien  
US-09-485-885-19

Query Match 25,9% Score 587; DB 4; Length 227;

Best Local Similarity 36.7%; Pred. No. 2.3e-51;  
Matches 152; Conservative 15; Mismatches 59; Indels 188; Gaps 9;

Qy 1 MDPSSSHSNMANTOMKSDKIIIAHAGSYLPHTLESKALAFQAQADYLEQDLAMTKDG 60  
|||||  
Db 1 MDPSSSHSNMANTOMKSDKIIIAHAGSYLPHTLESKALAFQAQADYLEQDLAMTKDG 60  
Qy 61 RLVIHDFHDLGLTDVAKKPPHRRKDGRIYVIDFTLKEIOSLEMTENFETMGKWSKSS 120  
|||||  
Db 61 RLVIHDFHDLGLTDVAKKPPHRRKDGRIYVIDFTLKEIOSLEMTENFETM----- 112  
Qy 121 VVGWPTVRMRRAEPAADGVGAASRDLEKKGAITSSNTAATNACAWLEAQEEVEGVFP 180  
|||||  
Db 113 -----AHGFKATQDILVH-----LEPQNE----- 133  
Qy 181 VTPQVPLRPMTYKAAVDLSHLEKKGLEGLHSORRODILDLYHTQGYFPMQNYTP 240  
|||||  
Db 134 ----IP-----VDLL-----GHQQLSDSEENDEID----- 155  
Qy 241 GPGVRYPLTFGCWYKLVPEPDKVEANKGENTSLLHPVSLHGMDDPEREVLWRPDSRL 300  
|||||  
Db 156 --GVNHQ----- 160  
Qy 301 AFHVARELHPE-YFKNCTSEPDPRLEPDKHPSQPKTACTNVCYKCCFHCQVCFTK 359  
|||||  
Db 161 --HLPARRAEPORHTMLCMCKCEARIELVVESSADDLRAFOQLNLTSFVCPWC----- 214  
Qy 360 ALGISYGRKKRRRRPPQGSQTHOVSLSKPTOSRSDPTGPKETSGHHHHH 413  
|||||  
Db 215 -----ASQTSQSGHHHHH 227

## RESULT 13

US-09-485-885-4  
; Sequence 4, Application US/09485885  
; Patent No. 6342224  
; GENERAL INFORMATION:  
; APPLICANT: Bruck, Claudine  
; APPLICANT: Cabezón Silva, Teresa  
; APPLICANT: Delisse, Anne-Marie Eva Fernande  
; APPLICANT: Gerard, Catherine Marie Ghislaine  
; APPLICANT: Lombardo-Bencheikh, Angela  
; TITLE OF INVENTION: Vaccine

FILE REFERENCE: B45107

; CURRENT APPLICATION NUMBER: US/09/485,885  
; CURRENT FILING DATE: 2000-02-18  
; PRIOR APPLICATION NUMBER: PCT/EP98/05285  
; PRIOR FILING DATE: 1998-08-17  
; PRIOR APPLICATION NUMBER: GB 9717953.5  
; PRIOR FILING DATE: 1997-08-22  
; NUMBER OF SEQ ID NOS: 23  
; SOFTWARE: FastSeq for Windows Version 3.0  
; SEQ ID NO 4  
; LENGTH: 273  
; TYPE: PRT  
; ORGANISM: Homo sapien  
US-09-485-885-4

Query Match 25.7% Score 582; DB 4; Length 273;

Best Local Similarity 36.3%; Pred. No. 9.7e-51;  
Matches 154; Conservative 24; Mismatches 84; Indels 162; Gaps 12;

Qy 1 MDPSSSHSNMANTOMKSDKIIIAHAGSYLPHTLESKALAFQAQADYLEQDLAMTKDG 60  
|||||  
Db 1 MDPSSSHSNMANTOMKSDKIIIAHAGSYLPHTLESKALAFQAQADYLEQDLAMTKDG 60  
Qy 61 RLVIHDFHDLGLTDVAKKPPHRRKDGRIYVIDFTLKEIOSLEMTENFETMGKWSKSS 120  
|||||  
Db 61 RLVIHDFHDLGLTDVAKKPPHRRKDGRIYVIDFTLKEIOSLEMTENFETM----- 112  
Qy 121 VVGWPTVRMRRAEPAADGVGAASRDLEKKGAITSSNTAATNACAWLEAQEEVEGVFP 180  
|||||  
Db 113 -AMQDQPERPKLPOLCTELQTTIHD-----ILECVYCKQ----- 149  
Qy 181 VTPQVPLRPMTYKAAVDLSHLEKKGLEGLHSORRODILDLYHTQ----- 230  
Db 150 -----LLRREYDFA-----FRDLGIVYRDGNPYAVCDKCL 180  
Qy 231 -YEPDQNYTPGCVRYPLTFGCWYKLVPEPDKVEANKGENTSLLHPVSLHGMDDPER 289  
|||||  
Db 181 KFYKISEYR-----HYCVSLYGTLE--QQYNKPLCDLLIRINCQKPLCPPEE 227  
Qy 290 EVLEWRFSRLAFHVAARELHPEYFKNCTSEPDPRLEPDKHPSQPKTACTNVCYKCC 349  
|||||  
Db 228 K--QRHLDDKKQRFHNR-----W-----TGRCMSCC 253  
Qy 350 FHCQVCFTTKALGISYGRKKRRRRPPQGSQTHOVSLSKPTOSRSDPTGPKETSGHH 409  
|||||  
Db 254 -----RSSRTRE-----TQLTSGHH 269  
Qy 410 HHHH 413  
Db 270 HHHH 273

## RESULT 14

US-08-679-493A-74  
; Sequence 74, Application US/08679493A  
; Patent No. 6303295  
; GENERAL INFORMATION:  
; APPLICANT: Taylor, Ethan W.  
; TITLE OF INVENTION: SELENOPROTEINS, CODING SEQUENCES AND METHODS  
; FILE REFERENCE: 55-95  
; CURRENT APPLICATION NUMBER: US/08/679,493A  
; CURRENT FILING DATE: 1996-07-12  
; PRIOR APPLICATION NUMBER: 60/001203  
; PRIOR FILING DATE: 1995-07-14  
; PRIOR APPLICATION NUMBER: 60/003,112  
; PRIOR FILING DATE: 1995-09-01  
; NUMBER OF SEQ ID NOS: 216  
; SOFTWARE: Patentin Ver. 2.0  
; SEQ ID NO 74  
; LENGTH: 151  
; TYPE: PRT  
; ORGANISM: Human immunodeficiency virus type 1

FEATURE:  
NAME/KEY: VARIANT  
LOCATION: (1)..(151)  
OTHER INFORMATION: X is selenocysteine.  
US-08-679-493A-74

Query Match 22.0%; Score 498.5; DB 4; Length 151;  
Best Local Similarity 73.0%; Pred. No. 1.le-42;  
Matches 103; Conservative 4; Mismatches 7; Indels 27; Gaps 3;  
QY 112 MGKSKSSVGVGPTVRRMRRAEPAADGVGAASRDLEKHGAITSSNTAATNAACAWLEA 171  
DB 1 MGKSKSSVGVGPTVRRMRRAEPAADGVGAASRDLEKHGAITSSNTAATNAACAWLEA 60  
QY 172 QEEVEGVFPVTPQVPLRPMTYKAADVLSHFLKEKGG----- 207  
DB 61 QEEVEGVFPVTPQVPLRPMTYKAADVLSHFLKEKGGTGRANSLPTKTRYPXSVLDLPHTRL 120  
QY 208 LEGL--IHSORRQDI-LDLWI 225  
DB 121 LPXLAEHLTRARGIGISTDLWM 141

RESULT 15  
US-08-450-257-1  
Sequence 1, Application US/08450257  
Patent No. 5652122  
GENERAL INFORMATION:  
APPLICANT: FRANKEL, Alan  
APPLICANT: PABO, Carl  
APPLICANT: BARSOUM, James G.  
APPLICANT: FAWELL, Stephen E.  
APPLICANT: PEPINSKY, R. B.  
TITLE OF INVENTION: TAT-DERIVED TRANSPORT POLYPEPTIDES  
NUMBER OF SEQUENCES: 69  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: FISH & NEAVE  
STREET: 1251 Avenue of the Americas  
CITY: New York  
STATE: New York  
COUNTRY: USA  
ZIP: 10020  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: Patent In Release #1.0, Version #1.25  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/450,257  
FILING DATE: 25-MAY-1995  
CLASSIFICATION: 514  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: US 08/235,403  
FILING DATE: 28-APR-1994  
APPLICATION NUMBER: US 07/934,375  
FILING DATE: 21-AUG-1992  
APPLICATION NUMBER: US 07/098,766  
FILING DATE: 28-JUL-1993  
APPLICATION NUMBER: PCT/US93/07833  
FILING DATE: 19-AUG-1993  
APPLICATION NUMBER: US 07/454,450  
FILING DATE: 21-DEC-1989  
APPLICATION NUMBER: US 07/636,662  
FILING DATE: 02-JAN-1991  
APPLICATION NUMBER: US 08/158,015  
FILING DATE: 24-NOV-1993  
ATTORNEY/AGENT INFORMATION:  
NAME: Haley Jr., James F.  
REGISTRATION NUMBER: 27,794  
REFERENCE/DOCKET NUMBER: B170 CIP 2  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: (212) 596-9000

TELEFAX: (212) 596-9090  
TELEX: 14-8367  
INFORMATION FOR SEQ ID NO: 1:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 86 amino acids  
TYPE: amino acid  
TOPOLOGY: linear  
MOLECULE TYPE: protein  
ORIGINAL SOURCE:  
ORGANISM: human immunodeficiency virus  
STRAIN: type 1  
US-08-450-257-1  
Query Match 21.5%; Score 486; DB 1; Length 86;  
Best Local Similarity 100.0%; Pred. No. 8.9e-42;  
Matches 85; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
QY 320 EPVDPRLPEWKHPGSOPTACTNCYCKKCCFHCQVCFITKALGISYGRKKRRRRRPPQG 379  
DB 2 EPVDPRLPEWKHPGSOPTACTNCYCKKCCFHCQVCFITKALGISYGRKKRRRRRPPQG 61  
QY 380 SOTHQVSLSKQPTSQSRGDPGTGPK 404  
DB 62 SOTHQVSLSKQPTSQSRGDPGTGPK 86  
RESULT 16  
US-08-450-246-1  
Sequence 1, Application US/08450246  
Patent No. 5670617  
GENERAL INFORMATION:  
APPLICANT: FRANKEL, Alan  
APPLICANT: PABO, Carl  
APPLICANT: BARSOUM, James G.  
APPLICANT: FAWELL, Stephen E.  
APPLICANT: PEPINSKY, R. B.  
TITLE OF INVENTION: TAT-DERIVED TRANSPORT POLYPEPTIDES  
NUMBER OF SEQUENCES: 69  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: FISH & NEAVE  
STREET: 1251 Avenue of the Americas  
CITY: New York  
STATE: New York  
COUNTRY: USA  
ZIP: 10020  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: Patent In Release #1.0, Version #1.25  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/450,246  
FILING DATE: 25-MAY-1995  
CLASSIFICATION: 514  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: US 08/235,403  
FILING DATE: 28-APR-1994  
APPLICATION NUMBER: US 07/934,375  
FILING DATE: 21-AUG-1992  
APPLICATION NUMBER: US 07/098,766  
FILING DATE: 28-JUL-1993  
APPLICATION NUMBER: PCT/US93/07833  
FILING DATE: 19-AUG-1993  
APPLICATION NUMBER: US 07/454,450  
FILING DATE: 21-DEC-1989  
APPLICATION NUMBER: US 07/636,662  
FILING DATE: 02-JAN-1991  
APPLICATION NUMBER: US 08/158,015  
FILING DATE: 24-NOV-1993  
ATTORNEY/AGENT INFORMATION:  
NAME: Haley Jr., James F.  
REGISTRATION NUMBER: 27,794

```
;
; REFERENCE/DOCKET NUMBER: B170 CIP 2
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (212) 596-9000
; TELEFAX: (212) 596-9090
; TELEX: 14-8367
; INFORMATION FOR SEQ ID NO: 1:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 86 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: protein
; ORIGINAL SOURCE:
; ORGANISM: human immunodeficiency virus
; STRAIN: type 1
; US-08-450-246-1

Query Match 21.5% Score 486; DB 1; Length 86;
Best Local Similarity 100.0%; Pred. No. 8.9e-42;
Matches 85; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 320 EPVDPRLPEWHPGSPQKTACTNCYKCKCFHCQVCFITKALGISYGRKKRRRRPPQG 379
DB 2 EPVDPRLPEWHPGSPQKTACTNCYKCKCFHCQVCFITKALGISYGRKKRRRRPPQG 61

QY 380 SOTHQVSLSKQPTQSRSRGTGPKE 404
DB 62 SOTHQVSLSKQPTQSRSRGTGPKE 86

RESULT 17
US-08-450-098-1
; Sequence 1, Application US/08450098
; Patent No. 5674980
; GENERAL INFORMATION:
; APPLICANT: FRANKEL, Alan
; APPLICANT: PABO, Carl
; APPLICANT: BARSOUM, James G.
; APPLICANT: FAWELL, Stephen E.
; APPLICANT: PEPINSKY, R. B.
; TITLE OF INVENTION: TAT-DERIVED TRANSPORT POLYPEPTIDES
; NUMBER OF SEQUENCES: 69
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: FISH & NEAVE
; STREET: 1251 Avenue of the Americas
; CITY: New York
; STATE: New York
; COUNTRY: USA
; ZIP: 10020
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patent In Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/450,098
; FILING DATE: 25-MAY-1995
; CLASSIFICATION: 514
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/235,403
; FILING DATE: 28-APR-1994
; APPLICATION NUMBER: US 07/934,375
; FILING DATE: 21-AUG-1992
; APPLICATION NUMBER: US 07/098,766
; FILING DATE: 28-JUL-1993
; APPLICATION NUMBER: PCT/US93/07833
; FILING DATE: 19-AUG-1993
; APPLICATION NUMBER: US 07/454,450
; FILING DATE: 21-DEC-1989
; APPLICATION NUMBER: US 07/636,662
; FILING DATE: 02-JAN-1991
; APPLICATION NUMBER: US 08/158,015
; FILING DATE: 24-NOV-1993

;
; ATTORNEY/AGENT INFORMATION:
; NAME: Haley Jr., James F.
; REGISTRATION NUMBER: 27,794
; REFERENCE/DOCKET NUMBER: B170 CIP 2
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (212) 596-9000
; TELEFAX: (212) 596-9090
; TELEX: 14-8367
; INFORMATION FOR SEQ ID NO: 1:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 86 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: protein
; ORIGINAL SOURCE:
; ORGANISM: human immunodeficiency virus
; STRAIN: type 1
; US-08-450-098-1

Query Match 21.5% Score 486; DB 1; Length 86;
Best Local Similarity 100.0%; Pred. No. 8.9e-42;
Matches 85; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 320 EPVDPRLPEWHPGSPQKTACTNCYKCKCFHCQVCFITKALGISYGRKKRRRRPPQG 379
DB 2 EPVDPRLPEWHPGSPQKTACTNCYKCKCFHCQVCFITKALGISYGRKKRRRRPPQG 61

QY 380 SOTHQVSLSKQPTQSRSRGTGPKE 404
DB 62 SOTHQVSLSKQPTQSRSRGTGPKE 86

RESULT 18
US-08-451-233-1
; Sequence 1, Application US/08451233
; Patent No. 5747641
; GENERAL INFORMATION:
; APPLICANT: FRANKEL, Alan
; APPLICANT: PABO, Carl
; APPLICANT: BARSOUM, James G.
; APPLICANT: FAWELL, Stephen E.
; APPLICANT: PEPINSKY, R. B.
; TITLE OF INVENTION: TAT-DERIVED TRANSPORT POLYPEPTIDES
; NUMBER OF SEQUENCES: 69
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: FISH & NEAVE
; STREET: 1251 Avenue of the Americas
; CITY: New York
; STATE: New York
; COUNTRY: USA
; ZIP: 10020
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patent In Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/451,233
; FILING DATE: 25-MAY-1995
; CLASSIFICATION: 514
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/235,403
; FILING DATE: 28-APR-1994
; APPLICATION NUMBER: US 07/934,375
; FILING DATE: 21-AUG-1992
; APPLICATION NUMBER: US 07/098,766
; FILING DATE: 28-JUL-1993
; APPLICATION NUMBER: PCT/US93/07833
; FILING DATE: 19-AUG-1993
; APPLICATION NUMBER: US 07/454,450
; FILING DATE: 21-DEC-1989
; APPLICATION NUMBER: US 07/636,662
; FILING DATE: 02-JAN-1991
; APPLICATION NUMBER: US 08/158,015
; FILING DATE: 24-NOV-1993
```

;; FILING DATE: 02-JAN-1991  
;; APPLICATION NUMBER: US 08/158,015  
;; FILING DATE: 24-NOV-1993  
;; ATTORNEY/AGENT INFORMATION:  
;; NAME: Haley Jr., James F.  
;; REGISTRATION NUMBER: 27,794  
;; REFERENCE/DOCKET NUMBER: B170 CIP 2  
;; TELECOMMUNICATION INFORMATION:  
;; TELEPHONE: (212) 596-9000  
;; TELEFAX: (212) 596-9090  
;; TELEX: 14-8367  
;; INFORMATION FOR SEQ ID NO: 1:  
;; SEQUENCE CHARACTERISTICS:  
;; LENGTH: 86 amino acids  
;; TYPE: amino acid  
;; TOPOLOGY: linear  
;; MOLECULE TYPE: protein  
;; ORIGINAL SOURCE:  
;; ORGANISM: human immunodeficiency virus  
;; STRAIN: type 1  
US-08-451-233-1

Query Match 21.5%; Score 486; DB 1; Length 86;  
Best Local Similarity 100.0%; Pred. No. 8.9e-42;  
Matches 85; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 320 EPVDRLPEWHPGQPKTACTNCYCKCCFHCQVCFITKALGISYGRKKRRRRPPOG 379  
Db 2 EPVDRLPEWHPGQPKTACTNCYCKCCFHCQVCFITKALGISYGRKKRRRRPPOG 61  
QY 380 SOTHQVSLSKQPTSQSRGDPGPK 404  
Db 62 SOTHQVSLSKQPTSQSRGDPGPK 86

RESULT 19  
US-08-450-236-1  
; Sequence 1, Application US/08450236  
; Patent No. 5804604  
; GENERAL INFORMATION:  
; APPLICANT: FRANKEL, Alan  
; APPLICANT: PABO, Carl  
; APPLICANT: BARSOUM, James G.  
; APPLICANT: FAWELL, Stephen E.  
; APPLICANT: PEPINSKY, R. B.  
; TITLE OF INVENTION: TAT-DERIVED TRANSPORT POLYPEPTIDES  
; NUMBER OF SEQUENCES: 69  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: FISH & NEAVE  
; STREET: 1251 Avenue of the Americas  
; CITY: New York  
; STATE: New York  
; COUNTRY: USA  
; ZIP: 10020  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Floppy disk  
; COMPUTER: IBM PC compatible  
; OPERATING SYSTEM: PC-DOS/MS-DOS  
; SOFTWARE: PatentIn Release #1.0, Version #1.25  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/08/450,236  
; FILING DATE: 25-MAY-1995  
; CLASSIFICATION: 514  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: US 08/235,403  
; FILING DATE: 28-APR-1994  
; APPLICATION NUMBER: US 07/934,375  
; FILING DATE: 21-AUG-1992  
; APPLICATION NUMBER: US 07/098,766  
; FILING DATE: 28-JUL-1993  
; APPLICATION NUMBER: PCT/US93/07833  
; FILING DATE: 19-AUG-1993

;; APPLICATION NUMBER: US 07/454,450  
;; FILING DATE: 21-DEC-1989  
;; APPLICATION NUMBER: US 07/636,662  
;; FILING DATE: 02-JAN-1991  
;; APPLICATION NUMBER: US 08/158,015  
;; FILING DATE: 24-NOV-1993  
;; ATTORNEY/AGENT INFORMATION:  
;; NAME: Haley Jr., James F.  
;; REGISTRATION NUMBER: 27,794  
;; REFERENCE/DOCKET NUMBER: B170 CIP 2  
;; TELECOMMUNICATION INFORMATION:  
;; TELEPHONE: (212) 596-9000  
;; TELEFAX: (212) 596-9090  
;; TELEX: 14-8367  
;; INFORMATION FOR SEQ ID NO: 1:  
;; SEQUENCE CHARACTERISTICS:  
;; LENGTH: 86 amino acids  
;; TYPE: amino acid  
;; TOPOLOGY: linear  
;; MOLECULE TYPE: protein  
;; ORIGINAL SOURCE:  
;; ORGANISM: human immunodeficiency virus  
;; STRAIN: type 1  
US-08-450-236-1

Query Match 21.5%; Score 486; DB 1; Length 86;  
Best Local Similarity 100.0%; Pred. No. 8.9e-42;  
Matches 85; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 320 EPVDRLPEWHPGQPKTACTNCYCKCCFHCQVCFITKALGISYGRKKRRRRPPOG 379  
Db 2 EPVDRLPEWHPGQPKTACTNCYCKCCFHCQVCFITKALGISYGRKKRRRRPPOG 61  
QY 380 SOTHQVSLSKQPTSQSRGDPGPK 404  
Db 62 SOTHQVSLSKQPTSQSRGDPGPK 86

RESULT 20  
US-07-808-452-1  
; Sequence 1, Application US/07808452  
; Patent No. 6063612  
; GENERAL INFORMATION:  
; APPLICANT: Jayasena, Sumedha D.  
; APPLICANT: Johnston, Brian H.  
; TITLE OF INVENTION: Antiviral Reagents Based on  
; TITLE OF INVENTION: RNA-Binding Proteins  
; NUMBER OF SEQUENCES: 24  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: Law Offices of Peter Dehlinger  
; STREET: 350 Cambridge Avenue, Suite 300  
; CITY: Palo Alto  
; STATE: CA  
; COUNTRY: USA  
; ZIP: 94306  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Floppy disk  
; COMPUTER: IBM PC compatible  
; OPERATING SYSTEM: PC-DOS/MS-DOS  
; SOFTWARE: PatentIn Release #1.0, Version #1.25  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/07/808,452  
; FILING DATE: 19911213  
; CLASSIFICATION: 514  
; ATTORNEY/AGENT INFORMATION:  
; NAME: Fabian, Gary R.  
; REGISTRATION NUMBER: 33,875  
; REFERENCE/DOCKET NUMBER: 8255-008  
; TELECOMMUNICATION INFORMATION:  
; TELEPHONE: (415) 324-0880  
; TELEFAX: (415) 324-0960  
; INFORMATION FOR SEQ ID NO: 1:

```
;
; SEQUENCE CHARACTERISTICS:
;   LENGTH: 86 amino acids
;   TYPE: AMINO ACID
;   TOPOLOGY: linear
;   MOLECULE TYPE: protein
;   HYPOTHETICAL: NO
;   ANTI-SENSE: NO
;   ORIGINAL SOURCE:
;     INDIVIDUAL ISOLATE: the sequence of the TAT protein of
;     INDIVIDUAL ISOLATE: HIV-1
; US-07-808-452-1

Query Match      21.5%; Score 486; DB 3; Length 86;
Best Local Similarity 100.0%; Pred. No. 8.9e-42;
Matches 85; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 320 EPVDPRLPEWKHPGSPKTACTNCTCYCKKCCFHCQVCFITKALGISYGRKKRRRRPPQG 379
Db 2 EPVDPRLPEWKHPGSPKTACTNCTCYCKKCCFHCQVCFITKALGISYGRKKRRRRPPQG 61

QY 380 STHQVSLSKOPTSQSRGDPGPK 404
Db 62 STHQVSLSKOPTSQSRGDPGPK 86

RESULT 22
US-09-124-900-6
; Sequence 6, Application US/09124900
; Patent No. 6268484
; GENERAL INFORMATION:
; APPLICANT: KATINGER, Hermann
; APPLICANT: BUCHACHER, Andrea
; APPLICANT: ERNST, Wolfgang
; APPLICANT: BALLAUN, Claudia
; APPLICANT: PURTSCHER, Martin
; APPLICANT: TRKOLA, Alexandra
; APPLICANT: PEDL, Renate
; APPLICANT: SCHMATZ, Christine
; APPLICANT: KLIMA, Annelles
; APPLICANT: STEINDL, Franz
; APPLICANT: MUSTER, Thomas
; TITLE OF INVENTION: HIV-Vaccines
; FILE REFERENCE: 1939-112P
; CURRENT APPLICATION NUMBER: US/09/124, 900
; CURRENT FILING DATE: 1998-07-30
; PRIOR APPLICATION NUMBER: PCT/EP95/01481
; PRIOR FILING DATE: 1995-04-19
; NUMBER OF SEQ ID NOS: 11
; SOFTWARE: PatentIn version 3.0
; SEQ ID NO 6
; LENGTH: 86
; TYPE: PRT
; ORGANISM: Human immunodeficiency virus type 1
; US-09-124-900-6

Query Match      21.5%; Score 486; DB 4; Length 86;
Best Local Similarity 100.0%; Pred. No. 8.9e-42;
Matches 85; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 320 EPVDPRLPEWKHPGSPKTACTNCTCYCKKCCFHCQVCFITKALGISYGRKKRRRRPPQG 379
Db 2 EPVDPRLPEWKHPGSPKTACTNCTCYCKKCCFHCQVCFITKALGISYGRKKRRRRPPQG 61

QY 380 STHQVSLSKOPTSQSRGDPGPK 404
Db 62 STHQVSLSKOPTSQSRGDPGPK 86

RESULT 23
US-09-451-905-19
; Sequence 19, Application US/09451905
; Patent No. 6306613
; GENERAL INFORMATION:
; APPLICANT: Robert Z. Florkiewicz
; APPLICANT: Andrew Baird
; APPLICANT: Dale E. Warnock
; TITLE OF INVENTION: MODULATORS OF LEADERLESS PROTEIN EXPORT
; FILE REFERENCE: 200124.402C4
; CURRENT APPLICATION NUMBER: US/09/451.905
; CURRENT FILING DATE: 1999-12-01
; NUMBER OF SEQ ID NOS: 48
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 19
; LENGTH: 86
; TYPE: PRT
; ORGANISM: Homo sapien
; US-09-451-905-19

SEQUENCE CHARACTERISTICS:
;   LENGTH: 86 amino acids
;   TYPE: AMINO ACID
;   TOPOLOGY: linear
;   MOLECULE TYPE: protein
;   HYPOTHETICAL: NO
;   ANTI-SENSE: NO
;   ORIGINAL SOURCE:
;     INDIVIDUAL ISOLATE: the sequence of the TAT protein of
;     INDIVIDUAL ISOLATE: HIV-1
; US-07-808-452-1

Query Match      21.5%; Score 486; DB 3; Length 86;
Best Local Similarity 100.0%; Pred. No. 8.9e-42;
Matches 85; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 320 EPVDPRLPEWKHPGSPKTACTNCTCYCKKCCFHCQVCFITKALGISYGRKKRRRRPPQG 379
Db 2 EPVDPRLPEWKHPGSPKTACTNCTCYCKKCCFHCQVCFITKALGISYGRKKRRRRPPQG 61

QY 380 STHQVSLSKOPTSQSRGDPGPK 404
Db 62 STHQVSLSKOPTSQSRGDPGPK 86

RESULT 21
US-09-030-613-19
; Sequence 19, Application US/09030613
; Patent No. 6083706
; GENERAL INFORMATION:
; APPLICANT: Florkiewicz, Robert Z.
; APPLICANT: Baird, J. Andrew
; TITLE OF INVENTION: INHIBITORS OF LEADERLESS PROTEIN EXPORT
; NUMBER OF SEQUENCES: 36
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: SEED and BERRY LLP
; STREET: 6300 Columbia Center, 701 Fifth Avenue
; CITY: Seattle
; STATE: Washington
; COUNTRY: USA
; ZIP: 98104
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/030.613
; FILING DATE: 25-FEB-1998
; CLASSIFICATION:
; ATTORNEY/AGENT INFORMATION:
; NAME: No. 6083706tenburg Ph.D., Carol
; REGISTRATION NUMBER: 39,317
; REFERENCE/DOCKET NUMBER: 760100.418C1
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (206) 622-4900
; TELEFAX: (206) 682-6031
; INFORMATION FOR SEQ ID NO: 19:
; SEQUENCE CHARACTERISTICS:
;   LENGTH: 86 amino acids
;   TYPE: amino acid
;   STRANDEDNESS:
;   TOPOLOGY: linear
; US-09-030-613-19

Query Match      21.5%; Score 486; DB 3; Length 86;
Best Local Similarity 100.0%; Pred. No. 8.9e-42;
Matches 85; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 320 EPVDPRLPEWKHPGSPKTACTNCTCYCKKCCFHCQVCFITKALGISYGRKKRRRRPPQG 379
Db 2 EPVDPRLPEWKHPGSPKTACTNCTCYCKKCCFHCQVCFITKALGISYGRKKRRRRPPQG 61
```

Query Match 21.5%; Score 486; DB 4; Length 86;  
Best Local Similarity 100.0%; Pred. No. 8.9e-42;  
Matches 85; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 320 EPVDRLEPFWKHGSPQKPTACTNCYKCCFHCQVCFITKALGISYGRKKRRRRPQ 379  
DB 2 EPVDRLEPFWKHGSPQKPTACTNCYKCCFHCQVCFITKALGISYGRKKRRRRPQ 61

QY 380 SOTHVSLSKOPTSOSRGDPTGPK 404  
DB 62 SOTHVSLSKOPTSOSRGDPTGPK 86

RESULT 24

US-08-235-403-1  
Sequence 1, Application US/08235403  
Patent No. 6316003  
GENERAL INFORMATION:  
APPLICANT: FRANKEL, Alan  
APPLICANT: PABO, Carl  
APPLICANT: BARSOUM, James G.  
APPLICANT: FAWELL, Stephen E.  
APPLICANT: PEPINSKY, R. B.  
TITLE OF INVENTION: TAT-DERIVED TRANSPORT POLYPEPTIDES  
NUMBER OF SEQUENCES: 69  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: FISH & NEAVE  
STREET: 1251 Avenue of the Americas  
CITY: New York  
STATE: New York  
COUNTRY: USA  
ZIP: 10020

COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: Patent in Release #1.0, Version #1.25  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/235,403  
FILING DATE:  
CLASSIFICATION: 514  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: US 07/934,375  
FILING DATE: 21-AUG-1992  
APPLICATION NUMBER: US 07/098,766  
FILING DATE: 28-JUL-1993  
APPLICATION NUMBER: PCT/US93/07833  
FILING DATE: 19-AUG-1993  
APPLICATION NUMBER: US 07/454,450  
FILING DATE: 21-DEC-1989  
APPLICATION NUMBER: US 07/636,662  
FILING DATE: 02-JAN-1991  
APPLICATION NUMBER: US 08/158,015  
FILING DATE: 24-NOV-1993  
ATTORNEY/AGENT INFORMATION:  
NAME: Haley Jr., James F.  
REGISTRATION NUMBER: 27,794  
REFERENCE/DOCKET NUMBER: B170 CIP 2  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: (212) 596-9000  
TELEFAX: (212) 596-9090  
TELEX: 14-8367

INFORMATION FOR SEQ ID NO: 1:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 86 amino acids  
TYPE: amino acid  
TOPOLOGY: linear  
MOLECULE TYPE: protein  
ORIGINAL SOURCE:  
ORGANISM: human immunodeficiency virus  
STRAIN: type 1  
US-08-235-403-1

Query Match 21.5%; Score 486; DB 4; Length 86;  
Best Local Similarity 100.0%; Pred. No. 8.9e-42;  
Matches 85; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 320 EPVDRLEPFWKHGSPQKPTACTNCYKCCFHCQVCFITKALGISYGRKKRRRRPQ 379  
DB 2 EPVDRLEPFWKHGSPQKPTACTNCYKCCFHCQVCFITKALGISYGRKKRRRRPQ 61

QY 380 SOTHVSLSKOPTSOSRGDPTGPK 404  
DB 62 SOTHVSLSKOPTSOSRGDPTGPK 86

RESULT 25

PCT-US92-10770-1  
Sequence 1, Application PC/TUS9210770  
GENERAL INFORMATION:  
APPLICANT: Jayasena, Sumedha D.  
APPLICANT: Johnston, Brian H.  
TITLE OF INVENTION: Antiviral Reagents Based on  
NUMBER OF SEQUENCES: 24  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: SRI International  
STREET: 333 Ravenswood Avenue  
CITY: Menlo Park  
STATE: CA  
COUNTRY: USA  
ZIP: 94025

COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: Patent in Release #1.0, Version #1.25  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: PCT/US92/10770  
FILING DATE: 19921211  
CLASSIFICATION:  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: US 07/808,452  
FILING DATE: 13-DEC-1991  
ATTORNEY/AGENT INFORMATION:  
NAME: Fabian, Gary R.  
REGISTRATION NUMBER: 33,875  
REFERENCE/DOCKET NUMBER: P-2962  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: (415) 859-4550  
TELEFAX: (415) 859-3880  
INFORMATION FOR SEQ ID NO: 1:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 86 amino acids  
TYPE: AMINO ACID  
TOPOLOGY: linear  
MOLECULE TYPE: protein  
HYPOTHETICAL: NO  
ANTI-SENSE: NO  
ORIGINAL SOURCE:  
INDIVIDUAL ISOLATE: the sequence of the TAT protein of  
INDIVIDUAL ISOLATE: HIV-1  
PCT-US92-10770-1

Query Match 21.5%; Score 486; DB 5; Length 86;  
Best Local Similarity 100.0%; Pred. No. 8.9e-42;  
Matches 85; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 320 EPVDRLEPFWKHGSPQKPTACTNCYKCCFHCQVCFITKALGISYGRKKRRRRPQ 379  
DB 2 EPVDRLEPFWKHGSPQKPTACTNCYKCCFHCQVCFITKALGISYGRKKRRRRPQ 61

QY 380 SOTHVSLSKOPTSOSRGDPTGPK 404

Db 62 SOTHQVSLSKQPTSQSRGDPGPK 86

## RESULT 26

PCR-US95-06077-2  
; Sequence 2, Application PC/TUS9506077  
; GENERAL INFORMATION:  
; APPLICANT: Immunobiology Research, Institute Inc.  
; TITLE OF INVENTION: Vaccine Interdiction of Extracellular  
; TITLE OF INVENTION: Transactivating Proteins of Human Immunodeficiency Virus  
; TITLE OF INVENTION: and Other Chronically Infecting Viruses Employing Similar  
; TITLE OF INVENTION: Intercellular Transactivating Strategies  
; NUMBER OF SEQUENCES: 38  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: Howson and Howson  
; STREET: Spring House Corporate Cntr, PO Box 457  
; CITY: Spring House  
; STATE: Pennsylvania  
; COUNTRY: USA  
; ZIP: 19477  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Floppy disk  
; COMPUTER: IBM PC compatible  
; OPERATING SYSTEM: PC-DOS/MS-DOS  
; SOFTWARE: Patentin Release #1.0, Version #1.25  
; APPLICATION NUMBER: PCT/US95/06077  
; FILING DATE:  
; CLASSIFICATION:  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: US 08/247,991  
; FILING DATE: 23-MAY-1994  
; ATTORNEY/AGENT INFORMATION:  
; NAME: Bak, Mary E.  
; REGISTRATION NUMBER: 31,215  
; REFERENCE/DOCKET NUMBER: IRI44PCT  
; TELECOMMUNICATION INFORMATION:  
; TELEPHONE: 215-540-9200  
; TELEFAX: 215-540-5818  
; INFORMATION FOR SEQ ID NO: 2:  
; SEQUENCE CHARACTERISTICS:  
; LENGTH: 86 amino acids  
; TYPE: amino acid  
; TOPOLOGY: linear  
; MOLECULE TYPE: protein  
PCT-US95-06077-2

Query Match 21.58; Score 486; DB 5; Length 86;  
Best Local Similarity 100.0%; Pred. No. 8.9e-42;  
Matches 85; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
Qy 320 EPVDPRLPEWHPGQSQTACTNCYKCCFCFQVCFITKALGISYGRKKRRRRPPQG 379  
Db 2 EPVDPRLPEWHPGQSQTACTNCYKCCFCFQVCFITKALGISYGRKKRRRRPPQG 61  
Qy 380 SOTHQVSLSKQPTSQSRGDPGPK 404  
Db 62 SOTHQVSLSKQPTSQSRGDPGPK 86

## RESULT 27

US-08-815-809-5  
; Sequence 5, Application US/08815809  
; Patent No. 6004777  
; GENERAL INFORMATION:  
; APPLICANT: TARTAGLIA, James  
; APPLICANT: GOEBEL, Scott J.  
; APPLICANT: COX, William I.  
; APPLICANT: GETTIG, Russell R.  
; APPLICANT: PINCUS, Steven E.  
; APPLICANT: PAOLETTI, Enzo  
; APPLICANT: JACOBS, Bertram L.

; TITLE OF INVENTION: VECTORS HAVING ENHANCED EXPRESSION, AND METHODS OF  
; FILE REFERENCE: 454310-3010  
; CURRENT APPLICATION NUMBER: US/08/815,809  
; CURRENT FILING DATE: 1997-03-12  
; NUMBER OF SEQ ID NOS: 23  
; SOFTWARE: Patentin Ver. 2.0  
; SEQ ID NO 5  
; LENGTH: 266  
; TYPE: PRT  
; ORGANISM: Vaccinia virus  
US-08-815-809-5

Query Match 21.18; Score 478.5; DB 3; Length 266;  
Best Local Similarity 77.3%; Pred. No. 2.7e-40;  
Matches 92; Conservative 4; Mismatches 6; Indels 17; Gaps 4;  
Qy 177 VGFPVTPQVPLRPMTYKAAVDLSHFLKEKGLGSLIHSQRRQDILDLWIYHTQGYFFDQW 236  
Db 136 VGFPVTPQVPLRPMTYKAAVDLSHFLKEKGLGSLIHSQRRQDILDLWIYHTQGYFFDQW 195  
Qy 237 NYTPGPGVRYPLTFGWCYKLVPEPDKEVEANKGENTSLHPVSLH-GMDDPPEVLEW 294  
Db 196 NYTPGPGVRYPLTFGWCYKLVPM-----IETV-----PVKLPQMDGP--KVKQW 238

## RESULT 28

US-08-816-155B-43  
; Sequence 43, Application US/08816155B  
; Patent No. 5990091  
; GENERAL INFORMATION:  
; APPLICANT: TARTAGLIA, JAMES  
; APPLICANT: COX, WILLIAM I.  
; APPLICANT: GETTIG, RUSSELL R.  
; APPLICANT: MARTINEZ, HECTOR  
; APPLICANT: PAOLETTI, ENZO  
; APPLICANT: PINCUS, STEVEN E.  
; TITLE OF INVENTION: VECTORS HAVING ENHANCED EXPRESSION, AND  
; NUMBER OF SEQUENCES: 48  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: FROMMER LAWRENCE & HAUG LLP  
; STREET: 745 FIFTH AVENUE  
; CITY: NEW YORK  
; STATE: NEW YORK  
; COUNTRY: USA  
; ZIP: 10151  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Floppy disk  
; COMPUTER: IBM PC compatible  
; OPERATING SYSTEM: PC-DOS/MS-DOS  
; SOFTWARE: Patentin Release #1.0, Version #1.30  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/08/816,155B  
; FILING DATE: 12-MAR-1997  
; CLASSIFICATION: 514  
; ATTORNEY/AGENT INFORMATION:  
; NAME: KOWALSKI, THOMAS J.  
; REGISTRATION NUMBER: 32,147  
; REFERENCE/DOCKET NUMBER: 454310-2990  
; TELECOMMUNICATION INFORMATION:  
; TELEPHONE: 212-588-0800  
; TELEFAX: 212-588-0500  
; INFORMATION FOR SEQ ID NO: 43:  
; SEQUENCE CHARACTERISTICS:  
; LENGTH: 280 amino acids  
; TYPE: amino acid  
; STRANDEDNESS: n/a  
; TOPOLOGY: linear  
; MOLECULE TYPE: amino acid  
US-08-816-155B-43



Query Match 21.1%; Score 478.5; DB 2; Length 280;  
 Best Local Similarity 77.3%; Pred. No. 2.9e-40;  
 Matches 92; Conservative 4; Mismatches 6; Indels 17; Gaps 4;

QY 177 VGFPVTPQVPLRPMTYKAADVLSHFLKEKGLGGLHSRRQDILDLDWIYHTQGYFPDQW 236  
 DB 150 VGFPVTPQVPLRPMTYKAADVLSHFLKEKGLGGLHSRRQDILDLDWIYHTQGYFPDQW 209

QY 237 NYTPGPGVRYPLTFGWCYKLVPM---IETV-----PVKLPKGMGDP--RVKQW 252

DB 210 NYTPGPGVRYPLTFGWCYKLVPM---IETV-----PVKLPKGMGDP--RVKQW 252

RESULT 29  
 US-09-079-587-43  
 ; Sequence 43, Application US/09079587  
 ; Patent No. 6130066  
 ; GENERAL INFORMATION:  
 ; APPLICANT: TARTAGLIA, JAMES  
 ; APPLICANT: COX, WILLIAM I.  
 ; APPLICANT: GETTIG, RUSSELL R.  
 ; APPLICANT: MARTINEZ, HECTOR  
 ; APPLICANT: PAOLETTI, ENZO  
 ; APPLICANT: PINCUS, STEVEN E.  
 ; TITLE OF INVENTION: VECTORS HAVING ENHANCED EXPRESSION, AND  
 ; METHODS OF MAKING AND USES THEREOF  
 ; NUMBER OF SEQUENCES: 48  
 ; CORRESPONDENCE ADDRESS:  
 ; ADDRESSEE: FROMMER LAWRENCE & HAUG LLP  
 ; STREET: 745 FIFTH AVENUE  
 ; CITY: NEW YORK  
 ; STATE: NEW YORK  
 ; COUNTRY: USA  
 ; ZIP: 10151  
 ; COMPUTER READABLE FORM:  
 ; MEDIUM TYPE: Floppy disk  
 ; COMPUTER: IBM PC compatible  
 ; OPERATING SYSTEM: PC-DOS/MS-DOS  
 ; SOFTWARE: PatentIn Release #1.0, Version #1.30  
 ; CURRENT APPLICATION DATA:  
 ; APPLICATION NUMBER: US/09/079.587  
 ; FILING DATE:  
 ; CLASSIFICATION:  
 ; PRIOR APPLICATION DATA:  
 ; APPLICATION NUMBER: US 08/916.155  
 ; FILING DATE: 12-NAR-1997  
 ; ATTORNEY/AGENT INFORMATION:  
 ; NAME: KOWALSKI, THOMAS J.  
 ; REGISTRATION NUMBER: 32,147  
 ; REFERENCE/DOCKET NUMBER: 454310-2990  
 ; TELECOMMUNICATION INFORMATION:  
 ; TELEPHONE: 212-588-0800  
 ; TELEFAX: 212-588-0500  
 ; INFORMATION FOR SEQ ID NO: 43:  
 ; SEQUENCE CHARACTERISTICS:  
 ; LENGTH: 280 amino acids  
 ; TYPE: amino acid  
 ; STRANDEDNESS: n/a  
 ; TOPOLOGY: linear  
 ; MOLECULE TYPE: amino acid  
 US-09-079-587-43

Query Match 21.1%; Score 478.5; DB 4; Length 280;  
 Best Local Similarity 77.3%; Pred. No. 2.9e-40;  
 Matches 92; Conservative 4; Mismatches 6; Indels 17; Gaps 4;

QY 177 VGFPVTPQVPLRPMTYKAADVLSHFLKEKGLGGLHSRRQDILDLDWIYHTQGYFPDQW 236  
 DB 150 VGFPVTPQVPLRPMTYKAADVLSHFLKEKGLGGLHSRRQDILDLDWIYHTQGYFPDQW 209

QY 237 NYTPGPGVRYPLTFGWCYKLVPM---IETV-----PVKLPKGMGDP--RVKQW 252

DB 210 NYTPGPGVRYPLTFGWCYKLVPM---IETV-----PVKLPKGMGDP--RVKQW 252

RESULT 30  
 US-08-505-210-1  
 ; Sequence 1, Application US/080505210  
 ; Patent No. 5981258  
 ; GENERAL INFORMATION:  
 ; APPLICANT: MEHTALI, Majid  
 ; APPLICANT: GUSS, Tania  
 ; TITLE OF INVENTION: COMPOSITION OF TRANS-DOMINANT VARIANTS  
 ; TITLE OF INVENTION: OF VIRAL PROTEINS FOR OBTAINING AN ANTIVIRAL EFFECT  
 ; NUMBER OF SEQUENCES: 7  
 ; CORRESPONDENCE ADDRESS:  
 ; ADDRESSEE: BURNS, DOANE, SWECKER & MATHIS, L.L.P.  
 ; STREET: 1737 King Street, Suite 500  
 ; CITY: Alexandria  
 ; STATE: Virginia  
 ; COUNTRY: United States  
 ; ZIP: 22314-2756  
 ; COMPUTER READABLE FORM:  
 ; MEDIUM TYPE: Floppy disk  
 ; COMPUTER: IBM PC compatible  
 ; OPERATING SYSTEM: PC-DOS/MS-DOS  
 ; SOFTWARE: PatentIn Release #1.0, Version #1.30  
 ; CURRENT APPLICATION DATA:  
 ; APPLICATION NUMBER: US/08/505,210  
 ; FILING DATE: 14-AUG-1995  
 ; CLASSIFICATION: 435  
 ; PRIOR APPLICATION DATA:  
 ; APPLICATION NUMBER: US 08/215,248  
 ; FILING DATE: 21-MAR-1994  
 ; PRIOR APPLICATION DATA:  
 ; APPLICATION NUMBER: WO PCT/FR94/01457  
 ; FILING DATE: 13-DEC-1994  
 ; ATTORNEY/AGENT INFORMATION:  
 ; NAME: Dadlo, Susan M.  
 ; REGISTRATION NUMBER: 40,373  
 ; REFERENCE/DOCKET NUMBER: 017753-066  
 ; TELECOMMUNICATION INFORMATION:  
 ; TELEPHONE: (703) 836-6620  
 ; TELEFAX: (703) 836-2021  
 ; INFORMATION FOR SEQ ID NO: 1:  
 ; SEQUENCE CHARACTERISTICS:  
 ; LENGTH: 86 amino acids  
 ; TYPE: amino acid  
 ; STRANDEDNESS: single  
 ; TOPOLOGY: linear  
 ; MOLECULE TYPE: protein  
 ; HYPOTHETICAL: NO  
 ; ANTI-SENSE: NO  
 ; ORIGINAL SOURCE:  
 ; ORGANISM: Human immunodeficiency virus type 1  
 ; STRAIN: Lai  
 ; INDIVIDUAL ISOLATE: sequence of the TAT protein of HIV-1  
 US-08-505-210-1

Query Match 20.8%; Score 470; DB 2; Length 86;  
 Best Local Similarity 96.5%; Pred. No. 3.7e-40;  
 Matches 82; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

QY 370 EPVDPRLPWPWKHPSQPKTACTNYCKKCCPHQVCFITTKALGISYGRKKRRRRPPQG 379  
 DB 2 EPVDPRLPWPWKHPSQPKTACTNYCKKCCPHQVCFITTKALGISYGRKKRRRRPPQG 61

QY 380 SOTHQVLSKOPTSQSRGDPGPK 404  
 DB 62 SOTHQVLSKOPTSQSRGDPGPK 86

Mon Aug 26 09:05:54 2002

us-09-509-239-21.ra1

Page 14

Search completed: August 26, 2002, 08:10:41  
Job time: 130 sec

GenCore version 4.5  
Copyright (c) 1993 - 2000 Compugen Ltd.

OM protein - protein search, using sw model

Run on: August 26, 2002, 08:15:42 ; Search time 40.39 Seconds  
(without alignments)  
395.919 Million cell updates/sec

Title: US-09-509-239-21

Perfect score: 2264

Sequence: 1 MDPSSSHSNWANTQMSDKI.....QSRGDPGPKETSGHHHHH 413

Scoring table:

BLOSUM62  
Gapop 10.0 , Gapext 0.5

Searched: 105224 seqs, 38719550 residues

Total number of hits satisfying chosen parameters: 105224

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : SwissProt\_40:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

# SUMMARIES

Result No.	Query %	Score	Match	Length	DB ID	Description
1	1130	49.9	206	1	NEF_HV1BR	P03406 human immun
2	1114	49.2	206	1	NEF_HV1PV	P03405 human immun
3	1111	49.1	206	1	NEF_HV1LW	Q70627 human immun
4	1108	48.9	206	1	NEF_HV1I2	P04324 human immun
5	1088.5	48.1	205	1	NEF_HV1B6	P05855 human immun
6	1019	45.0	218	1	NEF_HV1B1	P12479 human immun
7	1005.5	44.4	205	1	NEF_HV1S3	P19545 human immun
8	997.5	44.1	211	1	NEF_HV1OY	P20886 human immun
9	997	44.0	210	1	NEF_HV1A2	P03407 human immun
10	994	43.9	208	1	NEF_HV1RH	P05858 human immun
11	991.5	43.8	214	1	NEF_HV1Y2	P35959 human immun
12	981	43.3	216	1	NEF_HV1JR	P20867 human immun
13	974	43.0	208	1	NEF_HV1S1	P19546 human immun
14	927.5	41.0	206	1	NEF_HV1E1	P04604 human immun
15	911	40.2	212	1	NEF_HV1Z6	P04602 human immun
16	903.5	39.9	207	1	NEF_HV1ND	P18801 human immun
17	872	38.5	205	1	NEF_HV1U4	P24741 human immun
18	842	37.2	182	1	NEF_HV1M4	P05856 human immun
19	837	37.0	205	1	NEF_HV1Z8	P05859 human immun
20	834.5	36.9	209	1	NEF_HV1M2	P04603 human immun
21	832	36.7	205	1	NEF_HV1C2	P17664 chimpanzee
22	819	36.2	239	1	NEF_HV1SC	P05857 human immun
23	645	28.5	123	1	NEF_HV1H2	P04601 human immun
24	638	28.2	123	1	NEF_HV1B1	P03404 human immun
25	553.5	24.4	364	1	GLPQ_HAEIN	Q05282 haemophilus
26	486	21.5	86	1	TAT_HV1B1	P04606 human immun
27	481	21.2	86	1	TAT_HV1PV	P04607 human immun
28	478	21.1	86	1	TAT_HV1I2	P04326 human immun
29	470	20.8	86	1	TAT_HV1BR	P04610 human immun
30	461	20.4	102	1	TAT_HV1RH	P05908 human immun
31	454	20.1	86	1	TAT_HV1H2	P04608 human immun
32	446	19.7	101	1	TAT_HV1JR	P20879 human immun
33	439.5	19.4	97	1	NEF_HV1Z2	P12478 human immun

## ALIGNMENTS

RESULT 1

NEF\_HV1BR 1  
ID NEF\_HV1BR STANDARD; PRT; 206 AA.  
AC P03406;  
DT 21-JUL-1986 (Rel. 01, Created)  
DT 21-JUL-1986 (Rel. 01, Last sequence update)  
DT 16-OCT-2001 (Rel. 40, Last annotation update)  
DE Negative factor (F-protein) (27 kDa protein) (3'ORF).  
GN NEF.  
OS Human immunodeficiency virus type 1 (BRU isolate) (HIV-1), and  
OS Human immunodeficiency virus type 1 (New York-5 isolate) (HIV-1).  
OC Viruses; Retroviral viruses; Retroviridae; Lentivirus.  
OX NCBI\_TaxID=11686, 11698;  
RN [1]  
RP SEQUENCE FROM N.A. (CLONE PNL4-3).  
RC STRAIN=ISOLATE NEW YORK-5;  
RX MEDLINE=8509333; PubMed=2981635;  
RA Wain-Hobson S., Sonigo P., Danos O., Cole S., Alison M.;  
RT "Nucleotide sequence of the AIDS virus, LAV.";  
RL Cell 40:9-17(1985).  
RN [2]  
RP SEQUENCE FROM N.A. (CLONE PNL4-3).  
RC STRAIN=ISOLATE NEW YORK-5;  
RX Buckler C.E., Buckler-White A.J., Willey R.L., McCoy J.;  
RL Submitted (JUN-1988) to the EMBL/GenBank/DBJ databases.  
RN [3]  
RP POST-TRANSLATIONAL MODIFICATIONS, AND FUNCTION.  
RX MEDLINE=88039140; PubMed=3118220;  
RA Guy B., Kieny M.-P., Riviere Y., le Peuch C., Dott K., Girard M.,  
RA Montagnier L., Lecocq J.-P.;  
RT "HIV F3' orf encodes a phosphorylated GTP-binding protein resembling  
an oncogene product.";  
RL Nature 330:286-289(1987).  
RN [4]  
RP X-RAY CRYSTALLOGRAPHY (2.5 ANGSTROMS) OF 54-205 IN COMPLEX WITH NEF.  
RX MEDLINE=96279837; PubMed=8681387;  
RA Lee C.H., Sakseila K., Mirza U.A., Chait B.T., Kuriyan J.;  
RT "Crystal structure of the conserved core of HIV-1 Nef complexed with  
a Src family SH3 domain.";  
RL Cell 85:931-942(1996).  
CC -!- FUNCTION: NEF HAS GTPASE, GTP-BINDING AND AUTOPHOSPHORYLATING  
ACTIVITIES, IT SEEMS TO DOWN-REGULATE THE CD4(T4) ANTIGEN.  
CC -----  
CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
CC the European Bioinformatics Institute. There are no restrictions on its  
CC use by non-profit institutions as long as its content is in no way  
CC modified and this statement is not removed. Usage by and for commercial  
CC entities requires a license agreement (see <http://www.isb-sib.ch/announce/>  
CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
CC -----  
CC EMBL; K02013; AAB59752.1; -;  
DR EMBL; M19921; AAA44993.1; -;  
DR EMBL; A04321; CAA00353.1; -;  
DR PIR; A04008; ASLJFV  
DR PDB; 1EFN; 11-JAN-97.

34 439 19.4 253 1 NEF\_HV2KR Q74127 human immun  
35 437 19.3 101 1 TAT\_HV1SC P05906 human immun  
36 435 19.2 101 1 TAT\_HV1C4 P05907 human immun  
37 435 19.2 101 1 TAT\_HV1S1 P19553 human immun  
38 434 19.2 101 1 TAT\_HV1Y2 P35965 human immun  
39 431 19.0 101 1 TAT\_HV1MN P05905 human immun  
40 431 19.0 101 1 TAT\_HV1S3 P19552 human immun  
41 429.5 19.0 309 1 NEF\_SIVS4 P12482 simian immun  
42 426.5 18.8 261 1 NEF\_SIVSP P19501 simian immun  
43 421 18.6 257 1 NEF\_HV2D1 P17753 human immun  
44 419 18.5 101 1 TAT\_HV1A2 P04614 human immun  
45 419 18.5 101 1 TAT\_HV1OY P20893 human immun

```
DR HIV; K02013; NEFSBRU.
DR HIV; M19921; NEFSNL43.
DR InterPro; IPR001558; F-protein.
DR Pfam; PF00469; F-protein; 1.
KW AIDS; Myristate; GTP-binding; Phosphorylation; 3D-structure.
FT LIPID 2 MYRISTATE.
FT MOD_RES 15 15 PHOSPHORYLATION (BY PKC).
FT VARIANT 11 11 V -> I (IN CLONE PNL4-3).
FT VARIANT 15 15 T -> A (IN CLONE PNL4-3).
FT VARIANT 33 33 A -> V (IN CLONE PNL4-3).
FT VARIANT 51 51 T -> N (IN CLONE PNL4-3).
SQ SEQUENCE 206 AA; 23342 MW; 77453FC80B6004F2 CRC64;

Query Match 49.9%; Score 1130; DB 1; Length 206;
Best Local Similarity 100.0%; Pred. No. 1.4e-86;
Matches 206; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 112 MGGKWSKSSVVGWPTVRMRRAEPAADGVGAASRDLEKKGAISSNTAATNAACAWLEA 171
DB 1 MGGKWSKSSVVGWPTVRMRRAEPAADGVGAASRDLEKKGAISSNTAATNAACAWLEA 60

QY 172 QEEVEVGFPVTPQVPLRPMTYKAAVDLSHFLKEKGGLEGLHSQRQDILDWIYHTQY 231
DB 61 QEEVEVGFPVTPQVPLRPMTYKAAVDLSHFLKEKGGLEGLHSQRQDILDWIYHTQY 120

QY 232 FPDWQNTYTPGCVRYPLTFGWCYKLVPEPKVEANKGENTSLHHPVSLHGMDPPEV 291
DB 121 FPDWQNTYTPGCVRYPLTFGWCYKLVPEPKVEANKGENTSLHHPVSLHGMDPPEV 180

QY 292 LEWRFDSRLAFHHVARELHPEYFKNC 317
DB 181 LEWRFDSRLAFHHVARELHPEYFKNC 206

RESULT 2
NEF_HV1PV
ID NEF_HV1PV STANDARD; PRT; 206 AA.
AC P03405.
DT 21-JUL-1986 (Rel. 01, Created)
DT 21-JUL-1986 (Rel. 01, Last sequence update)
DT 16-OCT-2001 (Rel. 40, Last annotation update)
DE Negative factor (F-protein) (27 kDa protein) (3'ORF).
GN NEF.
OS Human immunodeficiency virus type 1 (PV22 isolate) (HIV-1).
OC Viruses; Retroviral viruses; Retroviridae; Lentivirus.
OX NCBI_TaxID=11700;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=85111157; PubMed=2982104;
RA Muesing M.A., Smith D.H., Cabradilla C.D., Benton C.V., Lasky L.A.,
RA Capon D.J.;
RT "Nucleic acid structure and expression of the human
RT AIDS/lymphadenopathy retrovirus.";
RL Nature 313:450-458(1985).
RN [2]
RP POST-TRANSLATIONAL MODIFICATIONS, AND FUNCTION.
RX MEDLINE=88039140; PubMed=3118220;
RA Guy B., Kieny M.-P., Riviere Y., le Peuch C., Dott K., Girard M.,
RA Montagnier L., Lecocq J.-P.;
RT "HIV F3' orf encodes a phosphorylated GTP-binding protein resembling
RT an oncogene product.";
RL Nature 330:266-269(1987).
CC -!- FUNCTION: NEF HAS GTPASE, GTP-BINDING AND AUTOPHOSPHORYLATING
CC ACTIVITIES, IT SEEMS TO DOWN-REGULATE THE CD4(T4) ANTIGEN.
CC
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC
```

```
CC EMBL; K02083; AAB59874.1;
DR EMBL; X01762; -, NOT_ANNOTATED_CDS.
DR PIR; A04007; ASLV.
DR HSP; P03406; IEFN.
DR HIV; K02083; NEFSVP22.
DR InterPro; IPR001558; F-protein.
DR Pfam; PF00469; F-protein; 1.
KW AIDS; Myristate; GTP-binding.
FT LIPID 2 MYRISTATE.
SQ SEQUENCE 206 AA; 23352 MW; EDE64281A17C6735 CRC64;

Query Match 49.2%; Score 1114; DB 1; Length 206;
Best Local Similarity 97.6%; Pred. No. 3e-85;
Matches 201; Conservative 3; Mismatches 2; Indels 0; Gaps 0;

QY 112 MGGKWSKSSVVGWPTVRMRRAEPAADGVGAASRDLEKKGAISSNTAATNAACAWLEA 171
DB 1 MGGKWSKSSVVGWPTVRMRRAEPAADGVGAASRDLEKKGAISSNTAATNAACAWLEA 60

QY 172 QEEVEVGFPVTPQVPLRPMTYKAAVDLSHFLKEKGGLEGLHSQRQDILDWIYHTQY 231
DB 61 QEEVEVGFPVTPQVPLRPMTYKAAVDLSHFLKEKGGLEGLHSQRQDILDWIYHTQY 120

QY 232 FPDWQNTYTPGCVRYPLTFGWCYKLVPEPKVEANKGENTSLHHPVSLHGMDPPEV 291
DB 121 FPDWQNTYTPGCVRYPLTFGWCYKLVPEPKVEANKGENTSLHHPVSLHGMDPPEV 180

QY 292 LEWRFDSRLAFHHVARELHPEYFKNC 317
DB 181 LEWRFDSRLAFHHVARELHPEYFKNC 206

RESULT 3
NEF_HV1LW
ID NEF_HV1LW STANDARD; PRT; 206 AA.
AC Q70627.
DT 15-JUL-1998 (Rel. 36, Created)
DT 15-JUL-1998 (Rel. 36, Last sequence update)
DT 16-OCT-2001 (Rel. 40, Last annotation update)
DE Negative factor (F-protein) (27 kDa protein) (3'ORF).
GN NEF.
OS Human immunodeficiency virus type 1 (LWI2.3 isolate) (HIV-1).
OC Viruses; Retroviral viruses; Retroviridae; Lentivirus.
OX NCBI_TaxID=82834;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=95127297; PubMed=7826699;
RA Reitz M.S. Jr., Hall L., Robert-Guroff M., Lautenberger J., Hahn B.M.,
RA Shaw G.M., Kong L.I., Weiss S.H., Waters D., Gallo R.C., Blattner W.;
RT "Viral variability and serum antibody response in a laboratory worker
RT infected with HIV type 1 (HTLV type IIIB).";
RL AIDS Res. Hum. Retroviruses 10:1143-1155(1994).
RN [2]
RP STRUCTURE BY NMR OF 56-206.
RX MEDLINE=9737445; PubMed=9194185;
RA Grzesiek S., Bax A., Hu J.S., Kaufman J., Palmer I., Stahl S.J.,
RA Tjandra N., Wingfield P.T.;
RT "Refined solution structure and backbone dynamics of HIV-1 Nef.";
RL Protein Sci. 6:1248-1263(1997).
CC -!- FUNCTION: NEF HAS GTPASE, GTP-BINDING AND AUTOPHOSPHORYLATING
CC ACTIVITIES, IT SEEMS TO DOWN-REGULATE THE CD4(T4) ANTIGEN.
CC
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC EMBL; U12055; AAA76691.1;
DR
```

PDB: 2NEF; 07-JUL-97.  
DR InterPro: IPR001558; F-protein.  
DR Pfam: PF00469; F-protein; 1.  
DR Pfam: PF00469; F-protein; 1.  
KW AIDS; Myristate: GTP-binding; 3D-structure.  
FT LIPID 2 MYRISTATE.  
SQ SEQUENCE 206 AA; 23414 MW; ED81F68F6B61278E CRC64;

Query Match 49.18; Score 1111; DB 1; Length 206;  
Best Local Similarity 97.18; Pred. No. 5.4e-85;  
Matches 200; Conservative 5; Mismatches 1; Indels 0; Gaps 0;

Qy 112 MGGKWSKSVVGPVTRRMRAEPAADVGAAASRDLEKKGAISSNTAATNAACAWLEA 171  
Db 1 MGGKWSKSVVGPVTRRMRAEPAADVGAAASRDLEKKGAISSNTAATNAACAWLEA 60  
Qy 172 QEEVEGVPVTPQVPLRPMYKAAVDLSHFLKKEGGLIHSQRQDILDWIYHTQY 231  
Db 61 QEEVEGVPVTPQVPLRPMYKAAVDLSHFLKKEGGLIHSQRQDILDWIYHTQY 120  
Qy 232 FPDWQNYTPGGVRYPLTFGWCYKLVPEPKVEANKGENTSLHHPVSLHGMDDPERV 291  
Db 121 FPDWQNYTPGGVRYPLTFGWCYKLVPEPKVEANKGENTSLHHPVSLHGMDDPERV 180  
Qy 292 LEWRFDLSLAFHHVARELHPEYFKNC 317  
Db 181 LEWRFDLSLAFHHVARELHPEYFKNC 206

RESULT 4  
NEF\_HV112  
ID NEF\_HV112 STANDARD; PRT; 206 AA.  
AC P04324;  
DT 20-MAR-1987 (Rel. 04, Created)  
DT 13-AUG-1987 (Rel. 05, Last sequence update)  
DT 16-OCT-2001 (Rel. 40, Last annotation update)  
DE Negative factor (F-protein) (27 kDa protein) (3'ORF).  
GN NEF.  
OS Human immunodeficiency virus type 1 (clone 12) (HIV-1).  
OC Viruses; Retroviral viruses; Retroviridae; Lentivirus.  
OX NCBI\_TaxID=11679;  
RN (1)  
RP SEQUENCE FROM N.A.  
RX MEDLINE=86177573; PubMed=3008154;  
RA Arya S.K., Gallo R.C.;  
RT "Three novel genes of human T-lymphotropic virus type III: immune reactivity of their products with sera from acquired immune deficiency syndrome patients."  
RT Proc. Natl. Acad. Sci. U.S.A. 83:2209-2213(1986).  
RN (2)  
RP POST-TRANSLATIONAL MODIFICATIONS, AND FUNCTION.  
RX MEDLINE=88039140; PubMed=3118220;  
RA Guy B., Kieny M.-P., Riviere Y., le Peuch C., Dott K., Girard M.,  
RA Montagnier L., Lecocq J.-P.;  
RT "HIV F/3' orf encodes a phosphorylated GTP-binding protein resembling an oncogene product."  
RT Nature 330:266-269(1987).  
RL Nature 330:266-269(1987).  
CC -!- FUNCTION: NEF HAS GTPASE, GTP-BINDING AND AUTOPHOSPHORYLATING ACTIVITIES, IT SEEMS TO DOWN-REGULATE THE CD4(T4) ANTIGEN.  
CC  
CC This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation - the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See <http://www.isb-sib.ch/announce/> or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
CC  
CC EMBL; M11840; AAA45001.1;  
DR PIR; A04006; ASLJ12.  
DR HSSP; P03406; 1EFN.  
DR HIV; M11840; NEFSPCV12.  
DR InterPro; IPR001558; F-protein.

Pfam: PF00469; F-protein; 1.  
KW AIDS; Myristate: GTP-binding.  
FT LIPID 2 MYRISTATE.  
SQ SEQUENCE 206 AA; 23366 MW; 218F5B2980F79A46 CRC64;

Query Match 48.98; Score 1108; DB 1; Length 206;  
Best Local Similarity 97.18; Pred. No. 9.5e-85;  
Matches 200; Conservative 4; Mismatches 2; Indels 0; Gaps 0;

Qy 112 MGGKWSKSVVGPVTRRMRAEPAADVGAAASRDLEKKGAISSNTAATNAACAWLEA 171  
Db 1 MGGKWSKSVVGPVTRRMRAEPAADVGAAASRDLEKKGAISSNTAATNAACAWLEA 60  
Qy 172 QEEVEGVPVTPQVPLRPMYKAAVDLSHFLKKEGGLIHSQRQDILDWIYHTQY 231  
Db 61 QEEVEGVPVTPQVPLRPMYKAAVDLSHFLKKEGGLIHSQRQDILDWIYHTQY 120  
Qy 232 FPDWQNYTPGGVRYPLTFGWCYKLVPEPKVEANKGENTSLHHPVSLHGMDDPERV 291  
Db 121 FPDWQNYTPGGVRYPLTFGWCYKLVPEPKVEANKGENTSLHHPVSLHGMDDPERV 180  
Qy 292 LEWRFDLSLAFHHVARELHPEYFKNC 317  
Db 181 LEWRFDLSLAFHHVARELHPEYFKNC 206

RESULT 5  
NEF\_HV1B8  
ID NEF\_HV1B8 STANDARD; PRT; 205 AA.  
AC P05855;  
DT 01-NOV-1998 (Rel. 09, Created)  
DT 01-NOV-1998 (Rel. 09, Last sequence update)  
DT 16-OCT-2001 (Rel. 40, Last annotation update)  
DE Negative factor (F-protein) (27 kDa protein) (3'ORF).  
GN NEF.  
OS Human immunodeficiency virus type 1 (BH8 isolate) (HIV-1).  
OC Viruses; Retroviral viruses; Retroviridae; Lentivirus.  
OX NCBI\_TaxID=11684;  
RN (1)  
RP SEQUENCE FROM N.A.  
RX MEDLINE=85111123; PubMed=2578615;  
RA Ratner L., Haseltine W., Patarca R., Livak K.J., Starcich B.R.,  
RA Josephs S.F., Doran E.R., Rafalski J.A., Whitehorn E.A.,  
RA Baumeister K., Ivanoff L., Petteway S.R. Jr., Pearson M.L.,  
RA Lautenberger J.A., Papas T.S., Grayeb J., Chang N.T., Gallo R.C.,  
RA Wong-Staal F.;  
RT "Complete nucleotide sequence of the AIDS virus, HTLV-III."  
RT Nature 313:277-284(1985).  
RN (2)  
RP POST-TRANSLATIONAL MODIFICATIONS, AND FUNCTION.  
RX MEDLINE=88039140; PubMed=3118220;  
RA Guy B., Kieny M.-P., Riviere Y., le Peuch C., Dott K., Girard M.,  
RA Montagnier L., Lecocq J.-P.;  
RT "HIV F/3' orf encodes a phosphorylated GTP-binding protein resembling an oncogene product."  
RT Nature 330:266-269(1987).  
RL Nature 330:266-269(1987).  
CC -!- FUNCTION: NEF HAS GTPASE, GTP-BINDING AND AUTOPHOSPHORYLATING ACTIVITIES, IT SEEMS TO DOWN-REGULATE THE CD4(T4) ANTIGEN.  
CC  
CC This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation - the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See <http://www.isb-sib.ch/announce/> or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
CC  
CC EMBL; K02011; -; NOT\_ANNOTATED\_CDS.  
DR HSSP; P03406; 1EFN.  
DR HIV; K02011; NEFSBH8.  
DR InterPro; IPR001558; F-protein.  
DR Pfam; PF00469; F-protein; 1.

KW AIDS; Myristate; GTP-binding.  
 FT LIPID 2 MYRISTATE.  
 SQ SEQUENCE 205 AA; 23305 MW; 8EC12F6650DD111 CRC64;

Query Match 48.1%; Score 1088.5; DB 1; Length 205;  
 Best Local Similarity 96.68; Pred. No. 3.9e-83;  
 Matches 199; Conservative 3; Mismatches 3; Indels 1; Gaps 1;  
 QY 112 MGGKWSKSSVVGWPTVRMRRAEPAADGGAASRDLEKKGGLGGLIHSORR 171  
 DB 1 MGGKWSKSSVVGWPAVRMRRAEPAADGGAASRDLEKKGGLGGLIHSORR 60  
 QY 172 QEEVEVGFPTQVPLRPMTYKAAVDLSHFLKKGGLGGLIHSORRQDILDWYHTQGY 231  
 DB 61 QEEVEVGFPTQVPLRPMTYKAAVDLSHFLKKGGLGGLIHSORRQDILDWYHTQGY 120  
 QY 232 FPDQNTYTPGVRYPLTFGCYKLVPEPKVEANKGENTSLHPVSLHGMDDPREV 291  
 DB 121 FPDQNTYTPGVRYPLTFGCYKLVPEPEK-EEANKGENTSLHPVSLHGMDDPREV 179  
 QY 292 LEWRFSRLAFHHVARELHPEYFKNC 317  
 DB 180 LEWRFSRLAFHHVARELHPEYFKNC 205

RESULT 6  
 NEF\_HV1BN  
 ID NEF\_HV1BN STANDARD; PRT; 218 AA.  
 AC P12479;  
 DT 01-OCT-1989 (Rel. 12, Created)  
 DT 01-OCT-1989 (Rel. 12, Last sequence update)  
 DT 16-OCT-2001 (Rel. 40, Last annotation update)  
 DE Negative factor (F-protein) (27 kDa protein) (3'ORF).  
 GN NEF.  
 OS Human immunodeficiency virus type 1 (1BR isolate) (HIV-1).  
 OC Viruses; Retroviral viruses; Retroviridae; Lentivirus.  
 OX NCBI\_TaxID=11693;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RX MEDLINE=89085613; PubMed=2789516;  
 RA Anand R., Thayer R., Srinivasan A., Nayyar S., Gardner M., Luciw P.,  
 RA Dandekar S.;  
 RT "Biological and molecular characterization of human immunodeficiency  
 RT virus (HIV-1BR) from the brain of a patient with progressive  
 RT dementia";  
 RL Virology 168:79-89(1989).  
 CC -1- FUNCTION: NEF HAS GTPASE, GTP-BINDING AND AUTOPHOSPHORYLATING  
 CC ACTIVITIES, IT SEEMS TO DOWN-REGULATE THE CD4(T4) ANTIGEN.  
 CC -1- MISCELLANEOUS: THIS VIRUS IS CYTOPATHICALLY ACTIVE AND WAS  
 CC HARVESTED FROM THE BRAIN TISSUE OF A NEUROLOGICAL AIDS PATIENT.  
 CC  
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
 CC the European Bioinformatics Institute. There are no restrictions on its  
 CC use by non-profit institutions as long as its content is in no way  
 CC modified and this statement is not removed. Usage by and for commercial  
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announcement/>  
 CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
 CC

EMBL; M21098; AAA4222.1;  
 DR PIR; D31687; ASLJBR.  
 DR HSP; P03406; IEFN.  
 DR HIV; M21098; NEFSBRVA.  
 DR InterPro; IPR001558; F-protein.  
 DR Pfam; PF00469; F-protein; 2.  
 KW AIDS; Myristate; GTP-binding.  
 FT LIPID 2 MYRISTATE (BY SIMILARITY).  
 SQ SEQUENCE 218 AA; 25032 MW; FC4DABA1045C460E CRC64;

Query Match 45.0%; Score 1019; DB 1; Length 218;  
 Best Local Similarity 84.1%; Pred. No. 2.4e-77;

Matches 185; Conservative 12; Mismatches 7; Indels 16; Gaps 2;  
 QY 112 MGGKWSKSSVVGWPTVRMR-----RAEPAADGGAASRDLEKKGGLGGLIHSORR 157  
 DB 1 MGGKWSK--MAGWSTVRMRMRRAEPAADGGAASRDLEKKGGLGGLIHSORR 58  
 QY 158 NTAATNACAWLEAEQEEVEVGFPTQVPLRPMTYKAAVDLSHFLKKGGLGGLIHSORR 217  
 DB 59 NTAATNACAWLEAEQEEVEVGFPTQVPLRPMTYKAAVDLSHFLKKGGLGGLIHSORR 118  
 QY 218 QDILDWYHTQGYFPDQNTYTPGVRYPLTFGCYKLVPEPKVEANKGENTSLH 277  
 DB 119 QDILDWYHTQGYFPDQNTYTPGVRYPLTFGCYKLVPEPEKIEANEENENSLH 178  
 QY 278 PVSUHGMDDPREVLEWRFSRLAFHHVARELHPEYFKNC 317  
 DB 179 PMSQHGMDDPREVLEWRFSRLAFHHVARELHPEYFKNC 218

RESULT 7  
 NEF\_HV1S3  
 ID NEF\_HV1S3 STANDARD; PRT; 205 AA.  
 AC P19545;  
 DT 01-FEB-1991 (Rel. 17, Created)  
 DT 01-FEB-1991 (Rel. 17, Last sequence update)  
 DT 16-OCT-2001 (Rel. 40, Last annotation update)  
 DE Negative factor (F-protein) (27 kDa protein) (3'ORF).  
 GN NEF.  
 OS Human immunodeficiency virus type 1 (SF33 isolate) (HIV-1).  
 OC Viruses; Retroviral viruses; Retroviridae; Lentivirus.  
 OX NCBI\_TaxID=11690;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RX MEDLINE=90317906; PubMed=2370688;  
 RA York-Higgins D., Cheng-Mayer C., Bauer D., Levy J.A., Dina D.;  
 RT "Human immunodeficiency virus type 1 cellular host range,  
 RT replication, and cytopathicity are linked to the envelope region of  
 RT the viral genome";  
 RL J. Virol. 64:4016-4020(1990).  
 CC -1- FUNCTION: NEF HAS GTPASE, GTP-BINDING AND AUTOPHOSPHORYLATING  
 CC ACTIVITIES, IT SEEMS TO DOWN-REGULATE THE CD4(T4) ANTIGEN.  
 CC  
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
 CC the European Bioinformatics Institute. There are no restrictions on its  
 CC use by non-profit institutions as long as its content is in no way  
 CC modified and this statement is not removed. Usage by and for commercial  
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announcement/>  
 CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
 CC  
 CC EMBL; M38427; AAA5068.1;  
 DR HSP; P03406; IEFN.  
 DR HIV; M38427; NEFSF33.  
 DR InterPro; IPR001558; F-protein.  
 DR Pfam; PF00469; F-protein; 1  
 KW AIDS; Myristate; GTP-binding.  
 FT LIPID 2 MYRISTATE (BY SIMILARITY).  
 SQ SEQUENCE 205 AA; 23318 MW; FFB419A1C5DFC9F3 CRC64;

Query Match 44.4%; Score 1005.5; DB 1; Length 205;  
 Best Local Similarity 86.4%; Pred. No. 2.9e-76;  
 Matches 178; Conservative 19; Mismatches 8; Indels 1; Gaps 1;

QY 112 MGGKWSKSSVVGWPTVRMRRAEPAADGGAASRDLEKKGGLGGLIHSORR 171  
 DB 1 MGGKWSKSK-MGWPAVRMRRAEPAADGGAASRDLEKKGGLGGLIHSORR 59  
 QY 172 QEEVEVGFPTQVPLRPMTYKAAVDLSHFLKKGGLGGLIHSORRQDILDWYHTQGY 231  
 DB 60 QEEVEVGFPTQVPLRPMTYKAAVDLSHFLKKGGLGGLIHSORRQDILDWYHTQGY 119  
 QY 232 FPDQNTYTPGVRYPLTFGCYKLVPEPKVEANKGENTSLHPVSLHGMDDPREV 291

DB 120 PDWQNYTPGVRPLTFGWCFLKVPVEPEKVEANEGENNSLLHPMSLHGMDPEKEV 179

QY 292 LEWFEDSLAFHHVARELHPEYKNC 317

DB 180 LKWKFDKSLAFRRHARELHPEYKDC 205

## RESULT 8

```
NEF_HV10Y ID NEF_HV10Y STANDARD; PRT; 211 AA.
AC P20886;
DT 01-FEB-1991 (Rel. 17, Last sequence update)
DT 01-FEB-1991 (Rel. 17, Last sequence update)
DT 16-OCT-2001 (Rel. 40, Last annotation update)
DE Negative factor (F-protein) (27 kDa protein) (3'ORF).
GN NEF.
OS Human immunodeficiency virus type 1 (OVI isolate) (HIV-1).
OC Viruses; Retrovirdae; Retroviridae; Lentivirus.
OX NCBI_TaxID=11699;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=90148544; PubMed=2559749;
RA Huet T., Dazza M.C., Brun-Vezinet F., Roelants G.E., Wain-Hobson S.;
RT "A highly defective HIV-1 strain isolated from a healthy Gabonese
RT individual presenting an atypical western blot.";
RL AIDS 3:707-715(1989).
CC -!- FUNCTION: NEF HAS GTPASE, GTP-BINDING AND AUTOPHOSPHORYLATING
CC ACTIVITIES, IT SEEMS TO DOWN-REGULATE THE CD4(T4) ANTIGEN.
CC -!- MISCELLANEOUS: THE OVI ISOLATE WAS TAKEN FROM THE BLOOD OF A
CC HEALTHY GABONESE INDIVIDUAL.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL; M26727; AAA83398.1;
DR HSSP; P03406; IEFN.
DR HIV; M26727; NEFSOYI.
DR InterPro; IPR001558; F-protein.
DR Pfam; PF00469; F-protein; 1.
KW AIDS; Myristate; GTP-binding.
FT LIPID 2
FT MYRISTATE.
SQ SEQUENCE 211 AA; 24067 MW; 64D846EEC55DAB8C CRC64;
```

Query Match 44.1%; Score 997.5; DB 1; Length 211;  
Best Local Similarity 82.9%; Pred. No. 1.4e-75;  
Matches 175; Conservative 22; Mismatches 9; Indels 5; Gaps 1;

QY 112 MGKWSKSSVGVWPTVRERMRRA----EPAADGVGAASRDLEKKGAISSNTAATNAAC 166

DB 1 MGKWSKSKMGKWPTRERDKRAELQPPPAAGVGGAASRDLEKKGAISSNTAATNADC 60

QY 167 AWLEAQEEVEGFPVTPQVPLRPMTYKAAVDLSHFLKEKGLGLIHSORQDILDWY 226

DB 61 AWLEAQEEVEGFPVTPQVPLRPMTYKAGDLSHFLKEKGLGLIYSOKRODILDWY 120

QY 227 HTQGYFPDQNTPTGPGVRYPLTFGCYKLVPEPKVEANKGENTSLHHPVSLHGMD 286

DB 121 HTQGYFPDQNTPTGPGVRYPLTFGCYKLVPEPKVEANKGENTSLHHPVSLHGMD 180

QY 287 PEREVLWFEDSLAFHHVARELHPEYKNC 317

DB 181 PEKVLWKFEDSLAFRRHARELHPEYKDC 211

## RESULT 9

NEF\_HV1A2 ID NEF\_HV1A2 STANDARD; PRT; 208 AA.

```
NEF_HV1A2 ID NEF_HV1A2 STANDARD; PRT; 210 AA.
AC P03407;
DT 21-JUL-1986 (Rel. 01, Created)
DT 21-JUL-1986 (Rel. 01, Last sequence update)
DT 16-OCT-2001 (Rel. 40, Last annotation update)
DE Negative factor (F-protein) (27 kDa protein) (3'ORF).
GN NEF.
OS Human immunodeficiency virus type 1 (ARV2/SF2 isolate) (HIV-1).
OC Viruses; Retrovirdae; Retroviridae; Lentivirus.
OX NCBI_TaxID=11685;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=85090453; PubMed=2578227;
RA Sanchez-Pescador R., Power M.D., Barr P.J., Steimer K.S.,
RA Stempien M.M., Brown-Shimer S.L., Gee W.W., Renard A., Randolph A.,
RA Levy J.A., Dina D., Luciw P.A.;
RT "Nucleotide sequence and expression of an AIDS-associated retrovirus
RT (ARV-2).";
RL Science 227:484-492(1985).
RN [2]
RP POST-TRANSLATIONAL MODIFICATIONS, AND FUNCTION.
RX MEDLINE=88039140; PubMed=3118220;
RA Guy B., Kieny M.-P., Riviere Y., le Peuch C., Dott K., Girard M.,
RA Montagnier L., Lecocq J.-P.;
RT "HIV F/3' orf encodes a phosphorylated GTP-binding protein resembling
RT an oncogene product.";
RL Nature 330:286-289(1987).
CC -!- FUNCTION: NEF HAS GTPASE, GTP-BINDING AND AUTOPHOSPHORYLATING
CC ACTIVITIES, IT SEEMS TO DOWN-REGULATE THE CD4(T4) ANTIGEN.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL; K02007; AAB59883.1;
DR PIR; A04009; ASLJO2.
DR HSSP; P03406; IEFN.
DR HIV; K02007; NEFSF2.
DR InterPro; IPR001558; F-protein.
DR Pfam; PF00469; F-protein; 1.
KW AIDS; Myristate; GTP-binding.
FT LIPID 2
FT MYRISTATE.
SQ SEQUENCE 210 AA; 24042 MW; ED255233F8A17DAB CRC64;
```

Query Match 44.0%; Score 997; DB 1; Length 210;  
Best Local Similarity 85.2%; Pred. No. 1.5e-75;  
Matches 179; Conservative 16; Mismatches 11; Indels 4; Gaps 1;

QY 112 MGKWSKSSVGVWPTVRERMR---RAEPAADGVGAASRDLEKKGAISSNTAATNAACA 167

DB 1 MGKWSKSKMGKWSAISREMRRAEPAADGVGAASRDLEKKGAISSNTAATNADCA 60

QY 168 WLEAQEEVEGFPVTPQVPLRPMTYKAAVDLSHFLKEKGLGLIHSORQDILDWY 227

DB 61 WLEAQEEVEGFPVTPQVPLRPMTYKAALDISHFLKEKGLGLIWSORQDILDWY 120

QY 228 TQGYFPDQNTPTGPGVRYPLTFGCYKLVPEPKVEANKGENTSLHHPVSLHGMD 287

DB 121 TQGYFPDQNTPTGPGVRYPLTFGCYKLVPEPKVEANKGENTSLHHPVSLHGMD 180

QY 288 PEREVLWFEDSLAFHHVARELHPEYKNC 317

DB 181 EKEVLWKFEDSLAFRRHARELHPEYKDC 210

## RESULT 10

NEF\_HV1RH ID NEF\_HV1RH STANDARD; PRT; 208 AA.

```

AC P05858;
DT 01-NOV-1988 (Rel. 09, Created)
DT 01-NOV-1988 (Rel. 09, Last sequence update)
DT 16-OCT-2001 (Rel. 40, Last annotation update)
DE Negative factor (F-protein) (27 kDa protein) (3'ORF).
GN NEF.
OS Human immunodeficiency virus type 1 (RE/HAT isolate) (HIV-1).
OC Viruses; Retrovirdae; Retroviridae; Lentivirus.
OX NCBI_TaxID=11701;
RN [1]
RP SEQUENCE FROM N.A.
RA Starcich B.R., Hahn B.H., Shaw G.M., McNeely P.D., Modrow S.,
RA Wolf H., Parks E.S., Parks W.P., Josephs S.F., Gallo R.C.,
RA Wong-Staal F.;
RL Submitted (XXX-1987) to the HIV data bank.
CC -1- FUNCTION: NEF HAS GTPASE, GTP-BINDING AND AUTOPHOSPHORYLATING
CC ACTIVITIES, IT SEEMS TO DOWN-REGULATE THE CD4(T4) ANTIGEN.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announcement/
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL; M17451; AAA45058.1;
DR HSP; P03406; IEFN.
DR HIV; M17451; NEFSRF.
DR InterPro; IPR001558; F-protein.
DR Pfam; PF00469; F-protein; 1.
DR AIDS; Myristate; GTP-binding.
FT LIPID 2
SQ SEQUENCE 208 AA; 23532 MW; 8F836FE8980F084C CRC64;

Query Match 43.9%; Score 994; DB 1; Length 208;
Best Local Similarity 84.6%; Pred. No. 2.7e-75;
Matches 176; Conservative 16; Mismatches 14; Indels 2; Gaps 1;

QY 112 MGGKWSKSSVGVWPTVRERMRRAEPAADGVAASRDLEKHGTAITSNTAA 171
DB 1 MGGKWSKSMGWPVAVRERMRRAEPAADGVAASRDLEKHGTAITSNTAA 171
QY 172 Q--EEEEVGFPVTPQVPLRPMTYKAAVDLSHFLKKGLEGLHSQRQDILDIWYHTQ 229
DB 61 QEEDEEVEGFPVTPQVPLRPMTYKAAVDLSHFLKKGLEGLHSQRQDILDIWYHTQ 120
QY 230 GYFPDQNYTPGQVRYPLTFGWCYKLVPEPKVEEANKGENTSLHPVSLHGMDPDER 289
DB 121 GYFPDQNYTPGQVRYPLTFGWCYKLVPEPKVEEANKGENTSLHPVSLHGMDPDER 180
QY 290 EVLEWRPDSRLAFHVAHRELHPEYKNC 317
DB 181 EYLVNKFDSRLAFHVAHRELHPEYKNC 208

RESULT 11
NEF_HV1Y2
ID NEF_HV1Y2 STANDARD; PRT; 214 AA.
AC P35959;
DT 01-JUN-1994 (Rel. 29, Created)
DT 01-JUN-1994 (Rel. 29, Last sequence update)
DT 16-OCT-2001 (Rel. 40, Last annotation update)
DE Negative factor (F-protein) (27 kDa protein) (3'ORF).
GN NEF.
OS Human immunodeficiency virus type 1 (YU-2 isolate) (HIV-1).
OC Viruses; Retrovirdae; Retroviridae; Lentivirus.
OX NCBI_TaxID=36377;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=93021387; PubMed=1404605;
RA Li Y., Hui H., Burgess C.J., Price R.W., Sharp P.M., Hahn B.H.,

RA Shaw G.M.;
RT "Complete nucleotide sequence, genome organization, and biological
RT properties of human immunodeficiency virus type 1 in vivo: evidence
RT for limited defectiveness and complementation.";
RL J. Virol. 66:6587-6600(1992).
CC -1- FUNCTION: NEF HAS GTPASE, GTP-BINDING AND AUTOPHOSPHORYLATING
CC ACTIVITIES, IT SEEMS TO DOWN-REGULATE THE CD4(T4) ANTIGEN.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announcement/
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL; M93258; -; NOT_ANNOTATED_CDS.
DR PIR; I44001; I44001.
DR HSP; P03406; IEFN.
DR InterPro; IPR001558; F-protein.
DR Pfam; PF00469; F-protein; 1.
DR AIDS; Myristate; GTP-binding.
FT LIPID 2
SQ SEQUENCE 214 AA; 24532 MW; E188D43D7B084D04 CRC64;

Query Match 43.8%; Score 991.5; DB 1; Length 214;
Best Local Similarity 84.2%; Pred. No. 4.5e-75;
Matches 181; Conservative 12; Mismatches 11; Indels 11; Gaps 2;

QY 112 MGGKWSKSSVGVWPTVR-----ERMRAEPAADGVAASRDLEKHGTAITSNTAA 161
DB 1 MGGKWSKSMAGWPTVRERMRRAEPAADGVAASRDLEKHGTAITSNTAA 60
QY 162 TNAACAWLEAQEEEEVEGFPVTPQVPLRPMTYKAAVDLSHFLKKGLEGLHSQRQDIL 221
DB 11 TNAACAWLEAQEEEEVEGFPVTPQVPLRPMTYKAAVDLSHFLKKGLEGLHSQRQDIL 120
QY 222 DLWYHTQGYFPDQNYTPGQVRYPLTFGWCYKLVPEPKVEEANKGENTSLHPVSL 281
DB 121 DLWYHTQGYFPDQNYTPGQVRYPLTFGWCYKLVPEPKVEEANKGENTSLHPVSL 179
QY 282 HGMDPDERLEWRPDSRLAFHVAHRELHPEYKNC 316
DB 180 HGMDPDERLEWRPDSRLAFHVAHRELHPEYKNC 214

RESULT 12
NEF_HV1JR
ID NEF_HV1JR STANDARD; PRT; 216 AA.
AC P20867;
DT 01-FEB-1991 (Rel. 17, Created)
DT 01-FEB-1991 (Rel. 17, Last sequence update)
DT 16-OCT-2001 (Rel. 40, Last annotation update)
DE Negative factor (F-protein) (27 kDa protein) (3'ORF).
GN NEF.
OS Human immunodeficiency virus type 1 (JRCFS isolate) (HIV-1).
OC Viruses; Retrovirdae; Retroviridae; Lentivirus.
OX NCBI_TaxID=11688;
RN [1]
RP SEQUENCE FROM N.A.
RA Koyanagi S., Chen I.S.-Y.;
RL Submitted (DEC-1988) to the HIV data bank.
CC -1- FUNCTION: NEF HAS GTPASE, GTP-BINDING AND AUTOPHOSPHORYLATING
CC ACTIVITIES, IT SEEMS TO DOWN-REGULATE THE CD4(T4) ANTIGEN.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announcement/
CC or send an email to license@isb-sib.ch).

```



```
CC -----
DR EMBL: M38429; AAB03750.1; -.
DR HSP: P03406; IEFN.
DR HIV: M38429; NEFSJRCFSF.
DR InterPro: IPR001558; F-protein.
DR Pfam: PF00469; F-protein; 1.
DR AIDS; Myristate; GTP-binding.
FT LIPID 2 2 MYRISTATE (BY SIMILARITY).
SQ SEQUENCE 216 AA; 24567 MW; D163FFA8C71529DC CRC64;

Query Match 43.3%; Score 981; DB 1; Length 216;
Best Local Similarity 81.9%; Pred. No. 3.4e-74;
Matches 177; Conservative 15; Mismatches 14; Indels 10; Gaps 1;

QY 112 MGGKWSKSSVGVPTVVRMRRAEPAAD-----GVGAASRDLEKHGAITSNTAA 161
DB 1 MGGKWSKSHVPGWSTVVRMRRAEPAATDRVQTEPAAVGVASRDLEKHGAITSNTAA 60
QY 162 TNAACAWLEAQEEVEGPPVTPQVPLRPMTYKAAVDLSHFLKEKGLEGLIHSQRQDIL 221
DB 61 TNADCAWLEAYEDEEVEGPPVTPQVPLRPMTYKAAVDLSHFLKEKGLEGLIYSQRQDIL 120
QY 222 DLWIYHTQGYFDPQNYTPGCVRYPLTFGWCYKLVPEPDKVEEANKGENTSLHPVSL 281
DB 121 DLWIYHTQGYFDPQNYTAGCVRYPLTFGWCYKLVPEPDKVEEANKGENTSLHPVSL 180
QY 282 HGMDDPERVLEWRDLSLAFHHVARELHPEYFKNC 317
DB 181 HGMDDPERVLEWRDLSLAFHHVARELHPEYFKNC 216

RESULT 13
NEF_HV1S1
ID NEF_HV1S1 STANDARD; PRT; 208 AA.
AC P19546;
DT 01-FEB-1991 (Rel. 17, Created)
DT 01-FEB-1991 (Rel. 17, Last sequence update)
DT 16-OCT-2001 (Rel. 40, Last annotation update)
DE Negative factor (F-protein) (27 kDa protein) (3'ORF).
GN NEF.
OS Human immunodeficiency virus type 1 (SF162 isolate) (HIV-1).
OC Viruses; Retroviral viruses; Retroviridae; Lentivirus.
OX NCBI_TaxID=11691;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=90347835; PubMed=2384920;
RA Cheng-Mayer C., Quiroga M., Tung J.W., Dina D., Levy J.;
RT "Viral determinants of human immunodeficiency virus type 1 T-cell or
RT macrophage tropism, cytopathogenicity, and CD4 antigen modulation.";
RL J. Virol. 64:4390-4398(1990).
CC -!- FUNCTION: NEF HAS GTPASE, GTP-BINDING AND AUTOPHOSPHORYLATING
CC ACTIVITIES, IT SEEMS TO DOWN-REGULATE THE CD4(T4) ANTIGEN.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL: M65024; AAA45073.1; -.
DR HSP: P03406; IEFN.
DR HIV: M38428; NEFSF162.
DR InterPro: IPR001558; F-protein.
DR Pfam: PF00469; F-protein; 1.
DR AIDS; Myristate; GTP-binding.
FT LIPID 2 2 MYRISTATE (BY SIMILARITY).
SQ SEQUENCE 208 AA; 23684 MW; A0B1007D14E46E32 CRC64;

Query Match 43.0%; Score 974; DB 1; Length 208;
Best Local Similarity 79.1%; Pred. No. 8.5e-70;
Matches 163; Conservative 24; Mismatches 18; Indels 1; Gaps 1;
```

```
Best Local Similarity 84.2%; Pred. No. 1.2e-73;
Matches 176; Conservative 16; Mismatches 13; Indels 4; Gaps 2;

QY 112 MGGKWSKSSVGVPTVVRMR---AEPADGVGAASRDLEKHGAITSNTAATNAACAW 168
DB 1 MGGKWSK-RMSGWSAVRERMRRAEPAADGVGAASRDLEKHGAITSNTAANADCAW 59
QY 169 LEAQEEVEGPPVTPQVPLRPMTYKAAVDLSHFLKEKGLEGLIHSQRQDILDWIYHT 228
DB 60 LEAQEDEDVGFVPRQVPLRPMTYKAAVDLSHFLKEKGLEGLIYSQRQDILDWIHT 119
QY 229 QGYFDPQNYTPGCVRYPLTFGWCYKLVPEPDKVEEANKGENTSLHPVSLHGMDDPE 288
DB 120 QGYFDPQNYTPGCVRYPLTFGWCYKLVPEPDKVEEANKGENTSLHPVSLHGMDDPE 179
QY 289 REVLEWRDLSLAFHHVARELHPEYFKNC 317
DB 180 KEVLVWRDLSLAFHHVARELHPEYFKNC 208

RESULT 14
NEF_HV1E1
ID NEF_HV1E1 STANDARD; PRT; 206 AA.
AC P04604;
DT 13-AUG-1987 (Rel. 05, Created)
DT 13-AUG-1987 (Rel. 05, Last sequence update)
DT 16-OCT-2001 (Rel. 40, Last annotation update)
DE Negative factor (F-protein) (27 kDa protein) (3'ORF).
GN NEF.
OS Human immunodeficiency virus type 1 (ELI isolate) (HIV-1).
OC Viruses; Retroviral viruses; Retroviridae; Lentivirus.
OX NCBI_TaxID=11689;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=86245056; PubMed=2424612;
RA Allison M., Wain-Hobson S., Montagnier L., Sonigo P.;
RT "Genetic variability of the AIDS virus: nucleotide sequence analysis
RT of two isolates from African patients.";
RL Cell 46:63-74(1986).
RN [2]
RP POST-TRANSLATIONAL MODIFICATIONS, AND FUNCTION.
RX MEDLINE=88039140; PubMed=3118220;
RA Guy B., Kieny M.-P., Riviere Y., le Peuch C., Dott K., Girard M.,
RA Montagnier L., Lecocq J.-P.;
RT "HIV F/3' orf encodes a phosphorylated GTP-binding protein resembling
RT an oncogene product.";
RL Nature 330:266-269(1987).
CC -!- FUNCTION: NEF HAS GTPASE, GTP-BINDING AND AUTOPHOSPHORYLATING
CC ACTIVITIES, IT SEEMS TO DOWN-REGULATE THE CD4(T4) ANTIGEN.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL: K03454; AAA44330.1; -.
DR EMBL: A07108; CAA00617.1; -.
DR HSP: P03406; IEFN.
DR HIV: K03454; NEFSCLI.
DR InterPro: IPR001558; F-protein.
DR Pfam: PF00469; F-protein; 1.
DR AIDS; Myristate; GTP-binding.
FT LIPID 2 2 MYRISTATE.
SQ SEQUENCE 206 AA; 23612 MW; 4CFF9F18A8AB503C CRC64;

Query Match 41.0%; Score 927.5; DB 1; Length 206;
Best Local Similarity 79.1%; Pred. No. 8.5e-70;
Matches 163; Conservative 24; Mismatches 18; Indels 1; Gaps 1;
```

```
QY 112 MGGKWSKSSVVGWPTVRERM-----RAEPAADGVGAASRDLEKKGAISSNTAATNAACAWLEA 171
DB 1 MGGKWSKSSIVGWPFAIRIRIRTPAADGVGAASRDLEKKGAISSNTASTNADCAWLEA 60
QY 172 QEE-EEVGFPVTPQVPLRPMTYKAAVDLSHFLKKEGGLGSLHSORQDIDLWIYHTQG 230
DB 61 QEESEVGFPPVPRQVPLRPMTYKALDLSHFLKKEGGLGSLHSORQDIDLWIYHTQG 120
QY 231 YPDQNTPTGPGVRYPLTFGWCYKLVPEPKVEEANKGENTSLHHPVSLHGMDDPERE 290
DB 121 IFPDQNTPTGPGIRYPLTFGWCYKLVPEPKVEEANKGENTSLHHPVSLHGMDDPERE 180
QY 291 VLEWREDSRLAFHVAHELHPEYKNC 316
DB 181 VLKWRFSRLAFEHKAREMHPEFYKNC 206
```

```
RESULT 15
NEF_HV126
ID NEF_HV126 STANDARD; PRT; 212 AA.
AC P04602;
DT 13-AUG-1987 (Rel. 05, Last sequence update)
DT 13-AUG-1987 (Rel. 05, Last sequence update)
DT 16-OCT-2001 (Rel. 40, Last annotation update)
DE Negative factor (F-protein) (27 kDa protein) (3'ORF).
GN NEF..
OS Human immunodeficiency virus type 1 (Zaire 6 isolate) (HIV-1).
OC Viruses; Retroviral viruses; Retroviridae; Lentivirus.
OX NCBI_TaxID=11708;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=87248097; PubMed=3036660;
RA Srinivasan A., Anand R., York D., Ranganathan P., Feorino P.,
RA Schochetman G., Curran J., Kalyanaram V.S., Luciw P.A.,
RA Sanchez-Pescador R.
RT "Molecular characterization of human immunodeficiency virus from
RT Zaire: nucleotide sequence analysis identifies conserved and variable
RT domains in the envelope gene."
RL Gene 52:71-82(1987).
RN [2]
RP POST-TRANSLATIONAL MODIFICATIONS, AND FUNCTION.
RX MEDLINE=88039140; PubMed=3118220;
RA Guy B., Kieny M.-P., Riviere Y., le Peuch C., Dott K., Girard M.,
RA Montagnier L., Lecocq J.-P.;
RT "HIV F3' orf encodes a phosphorylated GTP-binding protein resembling
RT an oncogene product."
RL Nature 330:266-269(1987).
CC -1- FUNCTION: NEF HAS GTPASE, GTP-BINDING AND AUTOPHOSPHORYLATING
CC ACTIVITIES. IT SEEMS TO DOWN-REGULATE THE CD4(T4) ANTIGEN.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (see http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
CC EMBL; M27323; AAA44874.1;
DR PIR; F26192; QQLJZR.
DR HSSP; P03406; 1EFN.
DR HIV; K03458; NEFSZ6.
DR InterPro; IPR001558; F-protein.
DR Pfam; PF00469; F-protein; 1.
DR AIDS; Myristate; GTP-binding.
FT LIPID 2
SQ SEQUENCE 212 AA; 24415 MW; EC095A1B263047CF CRC64;
```

```
Query Match 40.2%; Score 911; DB 1; Length 212;
Best Local Similarity 77.4%; Pred. No. 2.1e-68;
Matches 164; Conservative 23; Mismatches 19; Indels 6; Gaps 2;
```

```
QY 112 MGGKWSKSSVVGWPTVRERM-----RAEPAADGVGAASRDLEKKGAISSNTAATNAAC 166
DB 1 MGGKWSKSSIVGWPFAIRIRIRTPAADGVGAASRDLEKKGAISSNTRDTNADC 60
QY 167 AWLEAQEE-EEVGFPVTPQVPLRPMTYKAAVDLSHFLKKEGGLGSLHSORQDIDLWI 225
DB 61 AWLEAQEESEVGFPPVPRQVPLRPMTYKALDLSHFLKKEGGLGSLHSORQDIDLWI 120
QY 226 YHTQGYFPDQNTPTGPGVRYPLTFGWCYKLVPEPKVEEANKGENTSLHHPVSLHGM 285
DB 121 YHTQGYFPDQNTPTGPGIRYPLTFGWCYKLVPEPKVEEANKGENTSLHHPVSLHGM 180
QY 286 DPEREVLEWREDSRLAFHVAHELHPEYKNC 317
DB 181 DTEREVLKWRFSRLAFEHKAREMHPEFYKDC 212
```

```
RESULT 16
NEF_HVIND
ID NEF_HVIND STANDARD; PRT; 207 AA.
AC P18801;
DT 01-NOV-1990 (Rel. 16, Created)
DT 01-NOV-1990 (Rel. 16, Last sequence update)
DT 16-OCT-2001 (Rel. 40, Last annotation update)
DE Negative factor (F-protein) (27 kDa protein) (3'ORF).
GN NEF.
OS Human immunodeficiency virus type 1 (NDK isolate) (HIV-1).
OC Viruses; Retroviral viruses; Retroviridae; Lentivirus.
OX NCBI_TaxID=11695;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=90034200; PubMed=2806917;
RA Spire B., Sire J., Zachar V., Rey F., Barre-Sinoussi F., Galibert F.,
RA Hampe A., Chermann J.C.;
RT "Nucleotide sequence of HIV1-NDK: a highly cytopathic strain of the
RT human immunodeficiency virus."
RL Gene 81:275-284(1989).
CC -1- FUNCTION: NEF HAS GTPASE, GTP-BINDING AND AUTOPHOSPHORYLATING
CC ACTIVITIES. IT SEEMS TO DOWN-REGULATE THE CD4(T4) ANTIGEN.
CC -1- MISCELLANEOUS: NDK, ISOLATED FROM A ZAIRIAN PATIENT AFFECTED WITH
CC AIDS, AND IS A HIGHLY CYTOPATHOGENIC STRAIN.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (see http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
CC EMBL; M27323; AAA44874.1;
DR PIR; JQ0068; QQLJND.
DR HSSP; P03406; 1EFN.
DR HIV; M27323; NEFSNDK.
DR InterPro; IPR001558; F-protein.
DR Pfam; PF00469; F-protein; 1.
DR AIDS; Myristate; GTP-binding.
FT LIPID 2
SQ SEQUENCE 207 AA; 23748 MW; 09036C2F81D45D5E CRC64;
```

```
Query Match 39.9%; Score 903.5; DB 1; Length 207;
Best Local Similarity 77.3%; Pred. No. 8.4e-66;
Matches 160; Conservative 25; Mismatches 21; Indels 1; Gaps 1;
```

```
QY 112 MGGKWSKSSVVGWPTVRERMRAEPAADGVGAASRDLEKKGAISSNTAATNAACAWLEA 171
DB 1 MGGKWSKSSIVGWPFAIRIRIRTPAADGVGAASRDLEKKGAISSNTASTNADCAWLEA 60
QY 172 QEE-EEVGFPVTPQVPLRPMTYKAAVDLSHFLKKEGGLGSLHSORQDIDLWIYHTQG 230
DB 61 QEESEVGFPPVPRQVPLRPMTYKALDLSHFLKKEGGLGSLHSORQDIDLWIYHTQG 120
```

```
QY 231 YFPDQNTPGVYRPLTFGWCYKLVPEPKVEEANKGENTSLHPVSLHGDDPERE 290
Db 121 IFPDQNTPGGIRYPLTFGWCFLVPVDPQVEEATEREDNCLLHPMCOGMDPERQ 180
QY 291 VLEWRDSDRLAFHVAHELHPEYFKNC 317
Db 161 VLMWRFSRLALEKHARELHPEYFKDC 207

RESULT 17
NEF_HV1U4
ID NEF_HV1U4 STANDARD; PRT; 205 AA.
AC P24741;
DT 01-MAR-1992 (Rel. 21, Created)
DT 01-MAR-1992 (Rel. 21, Last sequence update)
DT 16-OCT-2001 (Rel. 40, Last annotation update)
DE Negative factor (F-protein) (27 kDa protein) (3'ORF).
GN NEF.
OS Human immunodeficiency virus type 1 (strain Ugandan / isolate U455)
OS (HIV-1).
OS Viruses; Retrovirdae; Retroviridae; Lentivirus.
OX NCBI_TaxID=11703;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=9109081; PubMed=2265025;
RA Oram J.D., Downing R.G., Roff M., Clegg J.C.S., Serwadda D.,
RA Carswell J.W.;
RT "Nucleotide sequence of a Ugandan HIV-1 provirus reveals genetic
RT diversity from other HIV-1 isolates.";
RL AIDS Res. Hum. Retroviruses 6:1073-1078 (1990).
CC -1- FUNCTION: NEF HAS GTPASE, GTP-BINDING AND AUTOPHOSPHORYLATING
CC ACTIVITIES, IT SEEMS TO DOWN-REGULATE THE CD4(T4) ANTIGEN.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (see http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
CC EMBL: M62320; AAA75023.1; -
CC HSP: P03406; IEFN.
CC InterPro: IPR001558; F-protein.
CC Pfam: PF00469; F-protein; 1.
CC AIDS; Myristate; GTP-binding.
CC LIPID 2
CC SEQUENCE 205 AA; 23253 MW; 869AB03E6E7893C4 CRC64;

Query Match 38.5%; Score 872; DB 1; Length 205;
Best Local Similarity 75.1%; Pred. No. 3.4e-65;
Matches 154; Conservative 24; Mismatches 27; Indels 0; Gaps 0;

QY 112 MGGKWSKSVVGVPTVRERMRRAEPAADGVAASRDLEKHGAISSNTAATNAACAWLEA 171
Db 1 MGGKWSKSRVPEVRKMRMTPTAAAGVAVSODLDKYGAVTSSTNSACAWLEA 60
QY 172 QEEVEVGFPVTPQVPLRPMYKAAVDLSHFLKEGGLGLIHSQRQDILDLYHTQGY 231
Db 61 QEEGVGFVPRVQVPLRPMYKAAVDLSHFLKEGGLGLIHSQRQDILDLYHTQGF 120
QY 232 FPDQNTPGVYRPLTFGWCYKLVPEPKVEEANKGENTSLHPVSLHGMDPERE 291
Db 121 FPDQNTPGVYRPLTFGWCYKLVPEPKVEEANKGENTSLHPVSLHGMDPERE 180
QY 292 LEWRDSDRLAFHVAHELHPEYFKNC 316
Db 181 LMKWFDSTLALKHAYELHPEYFKD 205

RESULT 18
NEF_HV1MN
```

```
ID NEF_HV1MN STANDARD; PRT; 182 AA.
AC P05856;
DT 01-NOV-1988 (Rel. 09, Created)
DT 01-NOV-1988 (Rel. 09, Last sequence update)
DT 16-OCT-2001 (Rel. 40, Last annotation update)
DE Negative factor (F-protein) (27 kDa protein) (3'ORF).
GN NEF.
OS Human immunodeficiency virus type 1 (MN isolate) (HIV-1).
OS Viruses; Retrovirdae; Retroviridae; Lentivirus.
OX NCBI_TaxID=11696;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=88219542; PubMed=3369091;
RA Gurgo C., Guo H.-G., Franchini G., Aldovini A., Collalti E.,
RA Farrell K., Wong-Staal F., Gallo R.C., Reitz M.S. Jr.;
RT "Envelope sequences of two new United States HIV-1 isolates.";
RL Virology 164:531-536 (1988).
CC -1- FUNCTION: NEF HAS GTPASE, GTP-BINDING AND AUTOPHOSPHORYLATING
CC ACTIVITIES, IT SEEMS TO DOWN-REGULATE THE CD4(T4) ANTIGEN.
CC -1- MISCELLANEOUS: THE MN ISOLATE WAS TAKEN FROM A PEDIATRIC AIDS
CC PATIENT IN 1984.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (see http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
CC EMBL: M17449; AAA44858.1; -
CC HSP: P03406; IEFN.
CC InterPro: IPR001558; F-protein.
CC Pfam: PF00469; F-protein; 1.
CC AIDS; Myristate; GTP-binding.
CC LIPID 2
CC SEQUENCE 182 AA; 20632 MW; 2C5B5A6E3EF13B26 CRC64;

Query Match 37.2%; Score 842; DB 1; Length 182;
Best Local Similarity 87.4%; Pred. No. 8.8e-63;
Matches 153; Conservative 11; Mismatches 7; Indels 4; Gaps 2;

QY 112 MGGKWSKSVVGVPTVRERMRRAEPAADGVAASRDLEKHGAISSNTAATNAACAW 168
Db 1 MGGKWSK-RVTGWPVTRERMRRAEPAADGVAASRDLEKHGAISSNTAATNAACAW 59
QY 169 LEAQEEVEVGFPVTPQVPLRPMYKAAVDLSHFLKEGGLGLIHSQRQDILDLYHT 228
Db 60 LEAQEEVEVGFPVTPQVPLRPMYKAAVDLSHFLKEGGLGLIHSQRQDILDLYHT 119
QY 229 QGYFPDQNTPGVYRPLTFGWCYKLVPEPKVEEANKGENTSLHPVSLHG 283
Db 120 QGYFPDQNTPGVYRPLTFGWCYKLVPEPKVEEANKGENTSLHPVSLHG 174

RESULT 19
NEF_HV1ZH
ID NEF_HV1ZH STANDARD; PRT; 205 AA.
AC P05859;
DT 01-NOV-1988 (Rel. 09, Created)
DT 01-NOV-1988 (Rel. 09, Last sequence update)
DT 16-OCT-2001 (Rel. 40, Last annotation update)
DE Negative factor (F-protein) (27 kDa protein) (3'ORF).
GN NEF.
OS Human immunodeficiency virus type 1 (Zaire H2321 isolate) (HIV-1).
OS Viruses; Retrovirdae; Retroviridae; Lentivirus.
OX NCBI_TaxID=11692;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=8922766; PubMed=2713163;
RA Srinivasan A., York D., Butler D. Jr., Jannoun-Nasr R., Getchell J.,
```

McCormick J., Ou C.Y., Myers G., Smith T., Chen E.;  
"Molecular characterization of HIV-1 isolated from a serum collected  
in 1976: nucleotide sequence comparison to recent isolates and  
generation of hybrid HIV-1;  
AIDS Res. Hum. Retroviruses 5:121-129(1989).  
-!- FUNCTION: NEF HAS GTPASE, GTP-BINDING AND AUTOPHOSPHORYLATING  
ACTIVITIES, IT SEEMS TO DOWN-REGULATE THE CD4(T4) ANTIGEN.  
-----  
This SWISS-PROT entry is copyright. It is produced through a collaboration  
between the Swiss Institute of Bioinformatics and the EMBL outstation -  
the European Bioinformatics Institute. There are no restrictions on its  
use by non-profit institutions as long as its content is in no way  
modified and this statement is not removed. Usage by and for commercial  
entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
-----  
EMBL; M15896; AAB53951.1; -;  
PIR; B44963; B44963.  
HSP; P03406; IEFN.  
HIV; M15896; NEFS2321.  
InterPro; IPR001558; F-protein.  
Pfam; PF00469; F-protein; 1.  
AIDS; Myristate; GTP-binding.  
FT LIPID 2 MYRISTATE (BY SIMILARITY).  
SQ SEQUENCE 205 AA; 23306 MW; EA55B18AF412A9D0 CRC64;  
-----  
Query Match 37.08; Score 837; DB 1; Length 205;  
Best Local Similarity 73.29; Pred. No. 2,7e-62;  
Matches 153; Conservative 23; Mismatches 25; Indels 8; Gaps 3;  
QY 112 MGGKWSKSSVVGWPTVRMRMR---AEPADGVGAASRDLEKHAITSNTAATNAACAW 168  
DB 1 MGNKWSK-----GWPAVRERIRIQTTPAPPAEGVGAASODLAKHAISSNTATNPDCAW 56  
QY 169 LEAQEE-EVGFPPVTPQVPLRPMTYKAAVDLSHFLKEKGLGEGLIHSORRODILDLWYH 227  
DB 57 LEAQEESEVGFPPVTPQVPLRPMTYKAAVDLSHFLKEKGLGEGLIHSORRODILDLWYH 116  
QY 228 TQGFPPDMQNTPGGVRYPTTFGCWYKLVPEVDKVEEANKGENTSLHVPVSLHGMDP 287  
DB 117 TQGFPPDMQNTPGGVRYPTTFGCWYKLVPEVDKVEEANKGENTSLHVPVSLHGMDP 176  
QY 288 EREVLEWRDLSRLAFHVAHRELHPEYKNC 316  
DB 177 EREVLEWRDLSRLAFHVAHRELHPEYKNC 205  
-----  
RESULT 20  
NEF\_HV1MA  
ID NEF\_HV1MA STANDARD; PRT; 209 AA.  
AC P04603;  
DT 13-AUG-1987 (Rel. 05, Created)  
DT 13-AUG-1987 (Rel. 05, Last sequence update)  
DT 16-OCT-2001 (Rel. 40, Last annotation update)  
DE Negative factor (F-protein) (27 kDa protein) (3'ORF).  
GN NEF.  
OS Human immunodeficiency virus type 1 (MAL isolate) (HIV-1).  
OC Viruses; Retrovirdae; Retroviridae; Lentivirus.  
OX NCBI\_TaxID=11697;  
RN [1]  
RP SEQUENCE FROM N.A.  
RX MEDLINE=86245056; PubMed=2424512;  
RA Alizon M., Wain-Hobson S., Montagnier L., Sonigo P.;  
RT "Genetic variability of the AIDS virus: nucleotide sequence analysis  
of two isolates from African patients.";  
RL Cell 46:63-74(1986).  
RN [2]  
RP POST-TRANSLATIONAL MODIFICATIONS, AND FUNCTION.  
RX MEDLINE=88039140; PubMed=3118220;  
RA Guy B., Kieny M.-P., Riviere Y., le Peuch C., Dott K., Girard M.,  
RA Montagnier L., Lecocq J.-P.;  
RT "HIV F/3' orf encodes a phosphorylated GTP-binding protein resembling

an oncogene product.";  
RL Nature 330:266-269(1987).  
CC -!- FUNCTION: NEF HAS GTPASE, GTP-BINDING AND AUTOPHOSPHORYLATING  
ACTIVITIES, IT SEEMS TO DOWN-REGULATE THE CD4(T4) ANTIGEN.  
-----  
This SWISS-PROT entry is copyright. It is produced through a collaboration  
between the Swiss Institute of Bioinformatics and the EMBL outstation -  
the European Bioinformatics Institute. There are no restrictions on its  
use by non-profit institutions as long as its content is in no way  
modified and this statement is not removed. Usage by and for commercial  
entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
-----  
EMBL; X04415; CAA28017.1; -;  
EMBL; A07116; CAA00624.1; -;  
HSP; P03406; IEFN.  
HIV; K03456; NEFSMAL.  
InterPro; IPR001558; F-protein.  
Pfam; PF00469; F-protein; 1.  
AIDS; Myristate; GTP-binding.  
FT LIPID 2 MYRISTATE.  
SQ SEQUENCE 209 AA; 23644 MW; DOB30A2442C8CC44 CRC64;  
-----  
Query Match 36.98; Score 834.5; DB 1; Length 209;  
Best Local Similarity 70.18; Pred. No. 4.4e-62;  
Matches 148; Conservative 26; Mismatches 30; Indels 7; Gaps 2;  
QY 112 MGGKWSKSSVVGWPTVRMRMRPAADGVG-----AASRDLEKHAITSNTAATNAAC 166  
DB 1 MGNKWSKSSVVGWPTVRMRMRPAADGVG-----AASRDLEKHAITSNTAATNAAC 60  
QY 167 AWLEAQEESEVGFPPVTPQVPLRPMTYKAAVDLSHFLKEKGLGEGLIHSORRODILDLWY 226  
DB 61 E-PPEEESEVGFPPVTPQVPLRPMTYKAAVDLSHFLKEKGLGEGLIHSORRODILDLWY 118  
QY 227 HTQGYFPDMQNTPGGVRYPTTFGCWYKLVPEVDKVEEANKGENTSLHVPVSLHGMD 286  
DB 119 HTQGYFPDMQNTPGGVRYPTTFGCWYKLVPEVDKVEEANKGENTSLHVPVSLHGMD 178  
QY 287 PEREVLEWRDLSRLAFHVAHRELHPEYKNC 317  
DB 179 AEREVLWKFDSSALRHAREQHPEYKNC 209  
-----  
RESULT 21  
NEF\_SIVCZ  
ID NEF\_SIVCZ STANDARD; PRT; 205 AA.  
AC P17664;  
DT 01-AUG-1990 (Rel. 15, Created)  
DT 01-AUG-1990 (Rel. 15, Last sequence update)  
DT 16-OCT-2001 (Rel. 40, Last annotation update)  
DE Negative factor (F-protein) (27 kDa protein) (3'ORF).  
GN NEF.  
OS Chimpanzee immunodeficiency virus (SIV(cpz)) (CIV).  
OC Viruses; Retrovirdae; Retroviridae; Lentivirus.  
OX NCBI\_TaxID=11723;  
RN [1]  
RP SEQUENCE FROM N.A.  
RX MEDLINE=90259077; PubMed=2188136;  
RA Huet T., Cheynier R., Meyerhans A., Roelants G., Wain-Hobson S.;  
RT "Genetic organization of a chimpanzee lentivirus related to HIV-1.";  
RL Nature 345:356-359(1990).  
CC -!- FUNCTION: NEF HAS GTPASE, GTP-BINDING AND AUTOPHOSPHORYLATING  
ACTIVITIES, IT SEEMS TO DOWN-REGULATE THE CD4(T4) ANTIGEN.  
-----  
This SWISS-PROT entry is copyright. It is produced through a collaboration  
between the Swiss Institute of Bioinformatics and the EMBL outstation -  
the European Bioinformatics Institute. There are no restrictions on its  
use by non-profit institutions as long as its content is in no way  
modified and this statement is not removed. Usage by and for commercial  
entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
-----

```

CC
DR EMBL; X52154; CAA36408.1; -.
DR DR PIR; S09991; ASLJTK.
DR DR HSSP; P03406; 1EFN.
DR DR HIV; X52154; NEFSCPZ.
DR DR InterPro; IPR001558; F-protein.
DR DR Pfam; PF00469; F-protein; 1.
DR DR AIDS; Myristate; GTP-binding.
DR KW AIDS; Myristate; GTP-binding.
DR FT LIPID 2
DR SQ SEQUENCE 205 AA; 23850 MW; 21EOA3EC99F1811F CRC64;

Query Match 36.7%; Score 832; DB 1; Length 205;
Best Local Similarity 70.2%; Pred. No. 6.9e-62;
Matches 144; Conservative 27; Mismatches 32; Indels 0; Gaps 0;

QY 112 MGGKWSKSSVGVGPTVERMRRAEPAADGVGAASRDLEKKGAISSNTAATNAACAWLEA 171
DB 1 MGGKWSKSSLVGVPEVRRRIREAPTAAGVGVEKSKDLERHGAITSRTPTNQTLAWLEE 60
: ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 172 QEEBEEVFPVTPQVPLRPMTYKAAVDLSHFLKEKGGLEGLIHSORRQDILDLYIHTQGY 231
DB 1 MGGKWSKSSLVGVPEVRRRIREAPTAAGVGVEKSKDLERHGAITSRTPTNQTLAWLEE 60
: ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
DB 61 MDNEEVGFPVRPQVTPRPMTYKAAFDLSHFLKEKGGLEGLVSRRRQEILDLYVYHTQGF 120
QY 232 FPDQONTPGGVRYPLTFGWCKYKLVPEVDKVEEANGKENTSLHVPYSLHGMDPEREV 291
DB 1 FPDQONTPTGTRFPLCFGWCFKVLPLTEQVEQANEQDNCLLHPICQHGMEDEKVEV 180
: ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 292 LEWRFDSRLAFHVAHELHPEYFKN 316
DB 1 LVRFPDSRLALRHIAEQHPEYK 205
: ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||

RESULT 22
NEF_HV1SC
ID NEF_HV1SC STANDARD; PRT; 239 AA.
AC P05857;
DT 01-NOV-1988 (Rel. 09, Created)
DT 01-NOV-1988 (Rel. 09, Last sequence update)
DT 16-OCT-2001 (Rel. 40, Last annotation update)
DE Negative factor (F-protein) (27 kDa protein) (3'ORF).
GN NEF.
OS Human immunodeficiency virus type 1 (SC isolate) (HIV-1).
OC Viruses; Retroid viruses; Retroviridae; Lentivirus.
OX NCBI_TaxID=11702;
RN (1)
RP MEDLINE=88219542; PubMed=3369091;
RA Gargo C., Guo H.-G., Franchini G., Aldovini A., Colliatti E.,
RA Farrell K., Wong-Staal F., Gallo R.C., Reitz M.S. Jr.;
RA "Envelope sequences of two new United States HIV-1 isolates.";
RL Virology 164:531-536(1988).
CC -1- FUNCTION: NEF HAS GTPASE, GTP-BINDING AND AUTOPHOSPHORYLATING
CC ACTIVITIES, IT SEEMS TO DOWN-REGULATE THE CD4(T4) ANTIGEN.
CC -1- MISCELLANEOUS: THE SC ISOLATE WAS TAKEN FROM AN ARC PATIENT IN
CC 1984 IN SOUTHERN CALIFORNIA.
CC
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (see http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC
CC EMBL; M17450; AAA45064.1; -.
CC HSSP; P03406; 1EFN.
CC HIV; M17450; NEF$SC.
CC InterPro; IPR001558; F-protein.
CC Pfam; PF00469; F-protein; 1.
CC AIDS; Myristate; GTP-binding.
CC FT LIPID 2
CC SQ SEQUENCE 239 AA; 26799 MW; 704A17E54763A99B CRC64;
MYRISTATE (BY SIMILARITY).

```

DR HIV: K03455; NEFSHXB2.  
DR InterPro: IPR001558; F-protein.  
DR Pfam: PF00469; F-protein; 1.  
KW AIDS; Myristate; GTP-binding; Phosphorylation.  
FT LIPID 2 MYRISTATE.  
FT MOD\_RES 15 15 PHOSPHORYLATION (BY PKC).  
SQ SEQUENCE 123 AA; 13692 MW; B5007753CCD244CF CRC64;

Query Match 28.5%; Score 645; DB 1; Length 123;  
Best Local Similarity 98.4%; Pred. No. 1.1e-46;  
Matches 121; Conservative 1; Mismatches 1; Indels 0; Gaps 0;  
QY 112 MGGKWSKSSVVGWPTVRRMRRAEPAADGVGAASRDLEKKGAISSNTAATNACAWLEA 171  
Db 1 MGGKWSKSSVVGWPTVRRMRRAEPAADGVGAASRDLEKKGAISSNTAATNACAWLEA 60  
QY 172 QEEVEGFPVTPQVPLRPMTYKAAVDLSHFLKKGGLGGLIHSORRODILDLYIHTQGY 231  
Db 61 QEEVEGFPVTPQVPLRPMTYKAAVDLSHFLKKGGLGGLIHSORRODILDLYIHTQGY 120  
QY 232 FPD 234  
Db 121 FPD 123

RESULT 24  
NEF\_HV1B1  
ID NEF\_HV1B1 STANDARD; PRT; 123 AA.  
AC P03404;  
DT 21-JUL-1986 (Rel. 01, Created)  
DT 13-AUG-1987 (Rel. 05, Last sequence update)  
DT 16-OCT-2001 (Rel. 40, Last annotation update)  
DE Negative factor (F-protein) (27 kDa protein) (3'ORF).  
GN NEF.  
OS Human immunodeficiency virus type 1 (BH10 isolate) (HIV-1).  
OC Viruses; Retroid viruses; Retroviridae; Lentivirus.  
OX NCBI\_TaxID=11678;  
RN [1]  
RP SEQUENCE FROM N.A.  
RX MEDLINE=85111123; PubMed=2578615;  
RA Ratner L., Haseltine W., Patarca R., Livak K.J., Starcich B.R.,  
RA Josephs S.F., Doran E.R., Rafalski J.A., Whitehorn E.A.,  
RA Baumeister K., Ivanoff L., Petteway S.R. Jr., Pearson M.L.,  
RA Lautenberger J.A., Papas T.S., Grayeb J., Chang N.T., Gallo R.C.,  
RA Wong-Staal F.;  
RT "Complete nucleotide sequence of the AIDS virus, HTLV-III.";  
RL Nature 313:277-284(1985).  
RN [2]  
RP POST-TRANSLATIONAL MODIFICATIONS, AND FUNCTION.  
RX MEDLINE=88039140; PubMed=3118220;  
RA Guy B., Kieny M.-P., Riviere Y., le Peuch C., Dott K., Girard M.,  
RA Montagnier L., Lecocq J.-P.;  
RT "HIV E/3" orf encodes a phosphorylated GTP-binding protein resembling  
RT an oncogene product.";  
RL Nature 330:266-269(1987).  
CC -1- FUNCTION: NEF HAS GTPASE, GTP-BINDING AND AUTOPHOSPHORYLATING  
CC ACTIVITIES, IT SEEMS TO DOWN-REGULATE THE CD4(T4) ANTIGEN.  
CC -1- MISCELLANEOUS: IN THIS ISOLATE A MUTATION IN POSITION 124 ADDS A  
CC STOP CODON AND THE PROTEIN IS MUCH SHORTER THAN IN OTHER ISOLATES  
CC (210 AA).  
CC -----  
CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
CC between the Swiss Institute of Bioinformatics and the EMBL outstation  
CC the European Bioinformatics Institute. There are no restrictions on its  
CC use by non-profit institutions as long as its content is in no way  
CC modified and this statement is not removed. Usage by and for commercial  
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
CC -----  
CC EMBL: M15654; AAA44206.1; -.  
CC PIR: A04005; ASLJH3.  
CC HSSP: Q70627; 2NEF.  
DR

DR HIV: M15654; NEFSBH102.  
DR InterPro: IPR001558; F-protein.  
DR Pfam: PF00469; F-protein; 1.  
KW AIDS; Myristate; GTP-binding.  
FT LIPID 2 MYRISTATE.  
SQ SEQUENCE 123 AA; 13606 MW; 0811735345F0EB8B CRC64;

Query Match 28.2%; Score 638; DB 1; Length 123;  
Best Local Similarity 97.8%; Pred. No. 4.2e-46;  
Matches 120; Conservative 0; Mismatches 3; Indels 0; Gaps 0;  
QY 112 MGGKWSKSSVVGWPTVRRMRRAEPAADGVGAASRDLEKKGAISSNTAATNACAWLEA 171  
Db 1 MGGKWSKSSVVGWPTVRRMRRAEPAADGVGAASRDLEKKGAISSNTAATNACAWLEA 60  
QY 172 QEEVEGFPVTPQVPLRPMTYKAAVDLSHFLKKGGLGGLIHSORRODILDLYIHTQGY 231  
Db 61 QEEVEGFPVTPQVPLRPMTYKAAVDLSHFLKKGGLGGLIHSORRODILDLYIHTQGY 120  
QY 232 FPD 234  
Db 121 FPD 123

RESULT 25  
GLPQ\_HAEIN  
ID GLPQ\_HAEIN STANDARD; PRT; 364 AA.  
AC Q06282;  
DT 01-NOV-1995 (Rel. 32, Created)  
DT 01-NOV-1995 (Rel. 32, Last sequence update)  
DT 16-OCT-2001 (Rel. 40, Last annotation update)  
DE Glycophosphoryl diester phosphodiesterase precursor (EC 3.1.4.46)  
DE (glycophosphodiester phosphodiesterase) (Surface-exposed lipoprotein  
DE D) (Protein D) (Immunoglobulin D-binding protein) (IGD-binding  
DE protein).  
GN GLPQ OR HPD OR HI0689.  
OS Haemophilus influenzae.  
OC Bacteria; Proteobacteria; gamma subdivision; Pasteurellaceae;  
OC Haemophilus.  
OX NCBI\_TaxID=727;  
RN [1]  
RP SEQUENCE FROM N.A.  
RX STRAIN=RD / KW20 / ATCC 51907;  
RX MEDLINE=95350630; PubMed=7542800;  
RA Fleischmann R.D., Adams M.D., White O., Clayton R.A., Kirkness E.F.,  
RA Kierlavage A.R., Bult C.J., Tomb J.-F., Dougherty B.A., Merrick J.M.,  
RA McKenney K., Sutton G., Fitzhugh W., Fields C.A., Gocayne J.D.,  
RA Scott J.D., Shirley R., Liu L.-I., Glodek A., Kelley J.M.,  
RA Weidman J.F., Phillips C.A., Spriggs T., Hedblom E., Cotton M.D.,  
RA Uterback T.R., Hanna M.C., Nguyen D.T., Saudek D.M., Brandon R.C.,  
RA Fine L.D., Fritchman J.L., Fuhrmann J.L., Geoghagen N.S.M.,  
RA Gnehm C.L., McDonald L.A., Small K.V., Fraser C.M., Smith H.O.,  
RA Venter J.C.;  
RT "Whole-genome random sequencing and assembly of Haemophilus  
RT influenzae Rd.";  
RL Science 269:496-512(1995).  
RN [2]  
RP SEQUENCE FROM N.A.  
RX STRAIN=NTHI 772;  
RX MEDLINE=91099948; PubMed=1987023;  
RA Janson H., Heden L.-O., Grubb A., Ruan M., Forsgren A.;  
RT "Protein D, an immunoglobulin D-binding protein of haemophilus  
RT influenzae: cloning, nucleotide sequence, and expression in  
RT Escherichia coli.";  
RL Infect. Immun. 59:119-125(1991).  
RN [3]  
RP SEQUENCE FROM N.A.  
RC STRAIN=MINNA / SEROTYPE B;  
RX MEDLINE=94011360; PubMed=8104899;  
RA Janson H., Ruan M., Forsgren A.;  
RT "Limited diversity of the protein D gene (hpd) among encapsulated and  
RT nonencapsulated Haemophilus influenzae strains.";









Best Local Similarity 98.8%; Pred. No. 4.7e-33;  
Matches 84; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 320 EPVDPRLPEPKHPSQPKTACTNCCYKCCFHCQVCFITKALGISYGRKKRRRPPQG 379  
|||||  
DB 2 EPVDPRLPEPKHPSQPKTACTNCCYKCCFHCQVCFITKALGISYGRKKRRRPPQG 61  
|||||

QY 380 SOTHOVSLSKOPTSQSRGDPGPK 404  
|||||  
DB 62 SOTHOVSLSKOPTSQSRGDPGPK 86  
|||||

## RESULT 29

TAT\_HV1BR  
ID TAT\_HV1BR STANDARD; PRT; 86 AA.

AC P04610; 1987 (Rel. 05, Created)

DT 13-AUG-1987 (Rel. 05, Last sequence update)

DT 13-AUG-1987 (Rel. 05, Last sequence update)

DT 16-OCT-2001 (Rel. 40, Last annotation update)

DE TAT protein (Transactivating regulatory protein).

GN TAT.

OS Human immunodeficiency virus type 1 (BRU isolate) (HIV-1).

OC Viruses; Retroviral viruses; Retroviridae; Lentivirus.

OX NCBI\_TaxID=11686;

RN [1]

RP SEQUENCE FROM N.A.

RX MEDLINE=8509333; PubMed=2981635;

RA Main-Hobson S., Sonigo P., Danos O., Cole S., Allison M.;

RT "Nucleotide sequence of the AIDS virus, LAV.,"

RL Cell 40:9-17(1985).

RN [2]

RP SEQUENCE FROM N.A. (CLONE PNL4-3).

RA Buckler C.E., Buckler-White A.J., Willey R.L., McCoy J.;

RL Submitted (JUN-1986) to the EMBL/GenBank/DBJ databases.

CC -1- FUNCTION: TRANSCRIPTIONAL REGULATOR THAT ACTS BY BINDING TO THE

CC TRANS-ACTIVATING RESPONSIVE SEQUENCE (TAR) RNA ELEMENT AND

CC ACTIVATES TRANSCRIPTION INITIATION AND/OR ELONGATION FROM THE LTR

CC PROMOTER.

CC -1- SUBUNIT: BINDS CYCLIN T1 (BY SIMILARITY).

CC -1- SUBCELLULAR LOCATION: NUCLEAR; NUCLEOLAR.

CC

CC This SWISS-PROT entry is copyright. It is produced through a collaboration

CC between the Swiss Institute of Bioinformatics and the EMBL outstation -

CC the European Bioinformatics Institute. There are no restrictions on its

CC use by non-profit institutions as long as its content is in no way

CC modified and this statement is not removed. Usage by and for commercial

CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>

CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).

CC

CC EMBL; K02013; AAB59745.1; -

DR EMBL; M19921; AAA4985.1; -

DR HIV; K02013; TAT\$BRU.

DR HIV; M19921; TAT\$NL43.

DR Pfam; PF00539; Tat; 1.

DR PRINTS; PR00055; HIVTATDOMAIN

DR Transcription regulation; Activator; RNA-binding; Nuclear protein;

KW AIDS.

FT VARIANT 24 24 T -> N (IN CLONE PNL4-3).

FT VARIANT 39 39 T -> M (IN CLONE PNL4-3).

FT VARIANT 58 61 PPG -> AHON (IN CLONE PNL4-3).

FT VARIANT 67 67 V -> A (IN CLONE PNL4-3).

FT VARIANT 77 77 P -> S (IN CLONE PNL4-3).

FT VARIANT 86 86 AA; 9769 MW; 9B1BA9A15FAF8A14 CRC64;

SQ SEQUENCE

## Query Match

Best Local Similarity 20.8%; Score 470; DB 1; Length 86;

Matches 82; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

QY 320 EPVDPRLPEPKHPSQPKTACTNCCYKCCFHCQVCFITKALGISYGRKKRRRPPQG 379  
|||||  
DB 2 EPVDPRLPEPKHPSQPKTACTNCCYKCCFHCQVCFITKALGISYGRKKRRRPPQG 61  
|||||

Search completed: August 26, 2002, 08:15:43

Job time: 347 sec

QY 380 SOTHOVSLSKOPTSQSRGDPGPK 404  
|||||  
DB 62 SOTHOVSLSKOPTSQSRGDPGPK 86  
|||||

## RESULT 30

TAT\_HV1BR  
ID TAT\_HV1BR STANDARD; PRT; 102 AA.

AC P05908;

DT 01-NOV-1988 (Rel. 09, Created)

DT 01-NOV-1988 (Rel. 09, Last sequence update)

DT 16-OCT-2001 (Rel. 40, Last annotation update)

DE TAT protein (Transactivating regulatory protein).

GN TAT.

OS Human immunodeficiency virus type 1 (RE/HAT isolate) (HIV-1).

OC Viruses; Retroviral viruses; Retroviridae; Lentivirus.

OX NCBI\_TaxID=11701;

RN [1]

RP SEQUENCE FROM N.A.

RX MEDLINE=86218077; PubMed=2423250;

RA Starcich B.R., Hahn B.H., Shaw G.M., McNeely P.D., Modrow S.,

RA Wolf H., Parks E.S., Parks W.P., Josephs S.F., Gallo R.C.,

RA Wong-Staal F.;

RT "Identification and characterization of conserved and variable

RT regions in the envelope gene of HTLV-III/LAV, the retrovirus of

RT AIDS.,"

RL Cell 45:637-648(1986).

CC -1- FUNCTION: TRANSCRIPTIONAL REGULATOR THAT ACTS BY BINDING TO THE

CC TRANS-ACTIVATING RESPONSIVE SEQUENCE (TAR) RNA ELEMENT AND

CC ACTIVATES TRANSCRIPTION INITIATION AND/OR ELONGATION FROM THE LTR

CC PROMOTER.

CC -1- SUBUNIT: BINDS CYCLIN T1 (BY SIMILARITY).

CC -1- SUBCELLULAR LOCATION: NUCLEAR; NUCLEOLAR.

CC

CC This SWISS-PROT entry is copyright. It is produced through a collaboration

CC between the Swiss Institute of Bioinformatics and the EMBL outstation -

CC the European Bioinformatics Institute. There are no restrictions on its

CC use by non-profit institutions as long as its content is in no way

CC modified and this statement is not removed. Usage by and for commercial

CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>

CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).

CC

CC EMBL; M17451; AAA45050.1; -

DR HIV; M17451; TAT\$SRF.

DR InterPro; IPR001831; HIV\_Tat.

DR Pfam; PF00539; Tat; 1.

DR PRINTS; PR00055; HIVTATDOMAIN.

DR Transcription regulation; Activator; RNA-binding; Nuclear protein;

KW AIDS.

SQ SEQUENCE 102 AA; 11538 MW; 3EACFB843C5195BD CRC64;

Query Match 20.4%; Score 461; DB 1; Length 102;

Best Local Similarity 91.9%; Pred. No. 1.5e-31;

Matches 79; Conservative 3; Mismatches 4; Indels 0; Gaps 0;

QY 320 EPVDPRLPEPKHPSQPKTACTNCCYKCCFHCQVCFITKALGISYGRKKRRRPPQG 379  
|||||  
DB 2 EPVDPRLPEPKHPSQPKTACTNCCYKCCFHCQVCFITKALGISYGRKKRRRPPQG 61  
|||||

QY 380 SOTHOVSLSKOPTSQSRGDPGPK 405  
|||||

DB 62 SOTHOVSLSKOPTSQSRGDPGPK 87  
|||||

GenCore version 4.5  
Copyright (c) 1993 - 2000 Compugen Ltd.

OM protein - protein search, using sw model

Run on: August 26, 2002, 08:14:56 ; Search time 86.53 Seconds  
(without alignments)  
825.689 Million cell updates/sec

Title: US-09-509-239-21  
Perfect score: 2264  
Sequence: 1 MDPSSSHSNMANTQMSDKI.....QSRGDTGPKETSGHHHHH 413

Scoring table:  
BLOSUM62  
Gapop 10.0 , Gapext 0.5

Searched: 562222 seqs, 172994929 residues

Total number of hits satisfying chosen parameters: 562222

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

Database :

1: sp\_archaea:\*  
2: sp\_bacteria:\*  
3: sp\_fungi:\*  
4: sp\_human:\*  
5: sp\_invertebrate:\*  
6: sp\_mammal:\*  
7: sp\_mhc:\*  
8: sp\_organelle:\*  
9: sp\_phage:\*  
10: sp\_plant:\*  
11: sp\_rodent:\*  
12: sp\_virus:\*  
13: sp\_vertebrate:\*  
14: sp\_unclassified:\*  
15: sp\_rvirus:\*  
16: sp\_bacteriap:\*  
17: sp\_archaeap:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

#### SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	1129	49.9	206	15 Q9WM24	Q9wm24 human immun
2	1124	49.6	206	15 Q9WLM4	Q9wlm4 human immun
3	1119	49.4	206	15 O40177	O40177 human immun
4	1115	49.2	206	15 Q90VU7	Q90vu7 human immun
5	1108	48.9	206	15 Q9PXW9	Q9pxw9 human immun
6	1105	48.8	206	15 Q9WM16	Q9wm16 human immun
7	1104	48.8	206	15 Q9WM30	Q9wm30 human immun
8	1101	48.6	206	15 Q9PXW8	Q9pxw8 human immun
9	1096	48.4	206	15 Q90179	Q90179 human immun
10	1096	48.4	206	15 Q85588	Q85588 aids-associ
11	1096	48.4	206	15 Q78244	Q78244 human immun
12	1090	48.2	206	15 Q9WLM7	Q9wlm7 human immun
13	1090	48.1	206	15 Q74905	Q74905 human immun
14	1087	48.0	206	15 Q89561	Q89561 human immun
15	1085	47.9	206	15 Q74913	Q74913 human immun
16	1080	47.7	206	15 Q74917	Q74917 human immun

17	1079	47.7	206	15 Q74914	Q74914 human immun
18	1079	47.7	206	15 Q74915	Q74915 human immun
19	1077	47.6	206	15 Q74909	Q74909 human immun
20	1076	47.5	202	15 Q9QPN3	Q9qpn3 human immun
21	1072	47.3	206	15 Q9Q596	Q9q596 human immun
22	1072	47.3	206	15 Q9Q595	Q9q595 human immun
23	1071	47.3	208	15 Q9VYU9	Q9vyu9 human immun
24	1068	47.2	206	15 Q9W7X3	Q9w7x3 human immun
25	1065	47.0	206	15 Q9DQ01	Q9dq01 human immun
26	1064	47.0	206	15 Q9W7U0	Q9w7u0 human immun
27	1062	46.9	206	15 Q93010	Q93010 human immun
28	1061	46.9	206	15 Q93012	Q93012 human immun
29	1060	46.8	206	15 Q9Q597	Q9q597 human immun
30	1057	46.7	206	15 Q9W7X2	Q9w7x2 human immun
31	1054	46.6	206	15 Q9Q594	Q9q594 human immun
32	1053	46.5	206	15 Q9Q5A3	Q9q5a3 human immun
33	1053	46.5	206	15 Q9DQ79	Q9dq79 human immun
34	1052	46.5	205	15 Q74823	Q74823 human immun
35	1051	46.4	206	15 Q74824	Q74824 human immun
36	1050	46.4	206	15 Q9QRD5	Q9qrd5 human immun
37	1050	46.4	206	15 Q9Q599	Q9q599 human immun
38	1049	46.3	206	15 Q89630	Q89630 human immun
39	1049	46.3	206	15 Q71956	Q71956 human immun
40	1048	46.3	206	15 Q9Q5A8	Q9q5a8 human immun
41	1048	46.3	206	15 Q72421	Q72421 human immun
42	1047	46.2	206	15 Q89537	Q89537 human immun
43	1047	46.2	206	15 Q9Q5D3	Q9q5d3 human immun
44	1047	46.2	206	15 Q72985	Q72985 human immun
45	1047	46.2	210	15 P90275	P90275 human immun

#### ALIGNMENTS

RESULT 1

Q9WM24 ID Q9WM24 PRELIMINARY; PRT; 206 AA.  
AC Q9WM24;  
DT 01-NOV-1999 (TREMBLrel. 12, Created)  
DT 01-NOV-1999 (TREMBLrel. 12, Last sequence update)  
DT 01-DEC-2001 (TREMBLrel. 19, Last annotation update)  
DE NEGATIVE FACTOR (F-PROTEIN) (27 KDA PROTEIN).  
GN NEF.  
OS Human immunodeficiency virus type 1.  
OC Viruses; Retroid viruses; Retroviridae; Lentivirus.  
OX NCBI\_TaxID=11676;  
RN [1]  
RP SEQUENCE FROM N.A.  
RC STRAIN=12-IT;  
RA Visco Comandini U., Yun Z., Paganelli R., Orlandi P., Salotti A.,  
RA Johansson B., Vahine A., Sonnerborg A.;  
RT "Hiv-1 nef mutations and clinical long-term non progression: a  
RT molecular epidemiology study";  
RL Submitted (JUN-1997) to the EMBL/GenBank/DBJ databases.  
CC -1- SIMILARITY: NEF HAS GTPASE, GTP-BINDING AND AUTOPHOSPHORYLATING  
CC ACTIVITIES, IT SEEMS TO DOWN-REGULATE THE CD4(T4) ANTIGEN.  
DR EMBL; AF011480; AAD01458.1;  
DR HSSP; P03406; 1EFN.  
DR InterPro; IPR001558; F-protein.  
DR Pfam; PF00469; F-protein; 1.  
KW AIDS; GTP-binding; Myristate.  
SQ SEQUENCE 206 AA; 23328 MW; FCC69458158F1A03 CRC64;

Query Match 49.9%; Score 1129; DB 15; Length 206;  
Best Local Similarity 99.5%; Pred. No. 3,6e-83;  
Matches 205; Conservative 1; Mismatches 0; Indels 0; Gaps 0;  
QY 112 MGGKWSKSVVGWPTVREMRRAEPAADGVGAASRDLEKHGAITSNTAATNAACAWLEA 171  
Db 1 MGGKWSKSVVGWPTVREMRRAEPAADGVGAASRDLEKHGAITSNTAATNAACAWLEA 60  
QY 172 QEEVEVGFPVTPQVPLRPMTYKAAVDLSHFLKEGKGLGGLIHSQRQDILDLIWIHTQCY 231

|||||  
 Db 61 QEEVEGFPVTPQVPLRPMTYKAAVDLSHFLKEKGLGSLHSQRRQDILDWYHTQGY 120  
 Qy 232 FPDQNTYTPGCVRYPLTFGWCYKLVPEPKVEEANKGENTSLHPVSLHGMDPDERV 291  
 Db 121 FPDQNTYTPGCVRYPLTFGWCYKLVPEPKVEEANKGENTSLHPVSLHGMDPDERV 180  
 Qy 292 LEWRDLSRLAFHHVARELHPYFKNC 317  
 Db 181 LEWRDLSRLAFHHVARELHPYFKNC 206

## RESULT 2

Q9WLM4  
 ID Q9WLM4 PRELIMINARY; PRT; 206 AA.  
 AC Q9WLM4  
 DT 01-NOV-1999 (TrEMBLrel. 12, Created)  
 DT 01-NOV-1999 (TrEMBLrel. 12, Last sequence update)  
 DT 01-DEC-2001 (TrEMBLrel. 19, Last annotation update)  
 DE NEGATIVE FACTOR (F-PROTEIN) (27 KDA PROTEIN).  
 GN NEF.  
 OS Human immunodeficiency virus type 1.  
 OC Viruses; Retrovirdae; Retroviridae; Lentivirus.  
 OX NCBI\_TaxID=11676;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN=21-SW;  
 RA Visco Comandini U., Yun Z., Paganelli R., Orlandi P., Salotti A.,  
 RA Johansson B., Wahine A., Sonnerborg A.;  
 RT "HIV-1 nef mutations and clinical long-term non progression: a  
 RT molecular epidemiology study";  
 RL Submitted (FEB-1998) to the EMBL/GenBank/DBJ databases.  
 CC -1- SIMILARITY: NEF HAS GTPASE, GTP-BINDING AND AUTOPHOSPHORYLATING  
 CC ACTIVITIES, IT SEEMS TO DOWN-REGULATE THE CD4 (T4) ANTIGEN.  
 DR EMBL; AF047087; AAD02461.1; -;  
 DR HSP; P03406; IEFN.  
 DR InterPro; IPR001558; F-protein.  
 DR Pfam; PF00469; F-protein; 1.  
 DR AIDS; GTP-binding; Myristate.  
 KW AIDS; GTP-binding; Myristate.  
 SQ SEQUENCE 206 AA; 23341 MW; FCD22B1CEB655BB9 CRC64;

Query Match 49.6%; Score 1124; DB 15; Length 206;  
 Best Local Similarity 99.0%; Pred. No. 9.2e-83;  
 Matches 204; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

Qy 112 MGGKWSKSSVVGWPTVRMRRAEPAADGVGAASRDLEKHAITSNTAATNAACAWLEA 171  
 Db 1 MGGKWSKSSVVGWPTVRMRRAEPAADGVGAASRDLEKHAITSNTAATNAACAWLEA 60  
 Qy 172 QEEVEGFPVTPQVPLRPMTYKAAVDLSHFLKEKGLGSLHSQRRQDILDWYHTQGY 231  
 Db 61 QEEVEGFPVTPQVPLRPMTYKAAVDLSHFLKEKGLGSLHSQRRQDILDWYHTQGY 120  
 Qy 232 FPDQNTYTPGCVRYPLTFGWCYKLVPEPKVEEANKGENTSLHPVSLHGMDPDERV 291  
 Db 121 FPDQNTYTPGCVRYPLTFGWCYKLVPEPKVEEANKGENTSLHPVSLHGMDPDERV 180  
 Qy 292 LEWRDLSRLAFHHVARELHPYFKNC 317  
 Db 181 LEWRDLSRLAFHHVARELHPYFKNC 206

## RESULT 3

Q40177  
 ID Q40177 PRELIMINARY; PRT; 206 AA.  
 AC Q40177  
 DT 01-JAN-1998 (TrEMBLrel. 05, Created)  
 DT 01-JAN-1998 (TrEMBLrel. 05, Last sequence update)  
 DT 01-DEC-2001 (TrEMBLrel. 19, Last annotation update)  
 DE NEGATIVE FACTOR (F-PROTEIN) (27 KDA PROTEIN).  
 GN NEF.  
 OS Human immunodeficiency virus type 1.

OC Viruses; Retrovirdae; Retroviridae; Lentivirus.  
 OX NCBI\_TaxID=11676;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN=PNL4-3;  
 RA Fang G., Weiser B., Visosky A., Burger H.;  
 RT "Constructing full-length chimeric HIV-1 molecular clones by PCR-  
 RT mediated recombination";  
 RL Submitted (MAY-1997) to the EMBL/GenBank/DBJ databases.  
 CC -1- SIMILARITY: NEF HAS GTPASE, GTP-BINDING AND AUTOPHOSPHORYLATING  
 CC ACTIVITIES, IT SEEMS TO DOWN-REGULATE THE CD4 (T4) ANTIGEN.  
 DR EMBL; AF003887; AAB64279.1; -;  
 DR HSP; P03406; IEFN.  
 DR InterPro; IPR001558; F-protein.  
 DR Pfam; PF00469; F-protein; 1.  
 DR AIDS; GTP-binding; Myristate.  
 SQ SEQUENCE 206 AA; 23428 MW; 9F94AEB9CAFD66F6 CRC64;

Query Match 49.4%; Score 1119; DB 15; Length 206;  
 Best Local Similarity 98.5%; Pred. No. 2.3e-82;  
 Matches 203; Conservative 1; Mismatches 2; Indels 0; Gaps 0;

Qy 112 MGGKWSKSSVVGWPTVRMRRAEPAADGVGAASRDLEKHAITSNTAATNAACAWLEA 171  
 Db 1 MGGKWSKSSVVGWPTVRMRRAEPAADGVGAASRDLEKHAITSNTAATNAACAWLEA 60  
 Qy 172 QEEVEGFPVTPQVPLRPMTYKAAVDLSHFLKEKGLGSLHSQRRQDILDWYHTQGY 231  
 Db 61 QEEVEGFPVTPQVPLRPMTYKAAVDLSHFLKEKGLGSLHSQRRQDILDWYHTQGY 120  
 Qy 232 FPDQNTYTPGCVRYPLTFGWCYKLVPEPKVEEANKGENTSLHPVSLHGMDPDERV 291  
 Db 121 FPDQNTYTPGCVRYPLTFGWCYKLVPEPKVEEANKGENTSLHPVSLHGMDPDERV 180  
 Qy 292 LEWRDLSRLAFHHVARELHPYFKNC 317  
 Db 181 LEWRDLSRLAFHHVARELHPYFKNC 206

## RESULT 4

Q90VU7  
 ID Q90VU7 PRELIMINARY; PRT; 206 AA.  
 AC Q90VU7  
 DT 01-DEC-2001 (TrEMBLrel. 19, Created)  
 DT 01-DEC-2001 (TrEMBLrel. 19, Last sequence update)  
 DT 01-DEC-2001 (TrEMBLrel. 19, Last annotation update)  
 DE NEF PROTEIN.  
 GN NEF.  
 OS Human immunodeficiency virus type 1.  
 OC Viruses; Retrovirdae; Retroviridae; Lentivirus.  
 OX NCBI\_TaxID=11676;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN=NL4-3;  
 RX MEDLINE=96036482; PubMed=7483282;  
 RA Salminen M.O., Koch C., Sanders-Buell E., Ehrenberg P.K.,  
 RA Michael N.L., Carr J.K., Burke D.S., McCutchan F.E.;  
 RT "Recovery of virtually full-length HIV-1 provirus of diverse subtypes  
 RT from primary virus cultures using the polymerase chain reaction.";  
 RL Virology 213:80-86(1995).  
 RN [2]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN=NL4-3;  
 RX MEDLINE=86281827; PubMed=3016298;  
 RA Adachi A., Gendelman H.E., Koenig S., Folks T., Willey R., Rabson A.,  
 RA Martin M.A.;  
 RT "Production of acquired immunodeficiency syndrome-associated  
 RT retrovirus in human and nonhuman cells transfected with an infectious  
 RT molecular clone.";  
 RL J. Virol. 59:284-291(1986).  
 DR EMBL; U26942; AAB60579.1; -;  
 SQ SEQUENCE 206 AA; 23367 MW; 65AF3B6184DC2FE7 CRC64;

```
Query Match          49.28; Score 1115; DB 15; Length 206;
Best Local Similarity 98.18; Pred. No. 4.9e-82;
Matches 202; Conservative 1; Mismatches 3; Indels 0; Gaps 0;

Qy 112 MGGKWSKSSVVGWPTVTRRMRRAEPAADGGAASRDLEKKGGLHSGRQDILDLWIYHTQY 171
Db 1 MGGKWSKSSVVGWPAVRERMRRAEPAADGGAASRDLEKKGGLHSGRQDILDLWIYHTQY 60

Qy 172 QEEVEGFPVTPQVPLRPMYKAAVDLSHFLKEKGGLEGLHSGRQDILDLWIYHTQY 231
Db 61 QEEVEGFPVTPQVPLRPMYKAAVDLSHFLKEKGGLEGLHSGRQDILDLWIYHTQY 120

Qy 232 FPDQNYTPGCGVRYPLTFGWCYKLVPEPKVEANKGENTSLHPVSLHGMDDPEREV 291
Db 121 FPDQNYTPGCGVRYPLTFGWCYKLVPEPKVEANKGENTSLHPVSLHGMDDPEREV 180

Qy 292 LEWRFDSRLAFHHVARELHPEYFKNC 317
Db 181 LEWRFDSRLAFHHVARELHPEYFKNC 206

RESULT 5
Q9PXM9 PRELIMINARY; PRT; 206 AA.
AC Q9PXM9;
DT 01-MAY-2000 (TrEMBLrel. 13, Created)
DT 01-MAY-2000 (TrEMBLrel. 13, Last sequence update)
DT 01-DEC-2001 (TrEMBLrel. 19, Last annotation update)
DE NEGATIVE FACTOR (F-PROTEIN) (27 KDA PROTEIN).
GN NEF.
OS Human immunodeficiency virus type 1.
OC Viruses; Retroid viruses; Retroviridae; Lentivirus.
OX NCBI_TaxID=11676;
RN [1]
RP SEQUENCE FROM N.A.
RC MEDLINE=93152025; PubMed=1301062;
RA Kohleisen B., Neumann M., Herrmann R., Brack-Werner R., Krohn K.J.,
RA Ovod V., Ranki A., Erfle V.;
RT "Cellular localization of Nef expressed in persistently HIV-1-infected
RT low-producer astrocytes.";
RL AIDS 6:1427-1436(1992).
CC -1- SIMILARITY: NEF HAS GTPASE, GTP-BINDING AND AUTOPHOSPHORYLATING
CC ACTIVITIES, IT SEEMS TO DOWN-REGULATE THE CD4(T4) ANTIGEN.
DR HSP; P03406; IEFN.
DR InterPro: IPR001558; F-protein.
DR Pfam: PF00469; F-protein; 1.
DR AIDS; GTP-binding; Myristate.
KW AIDS; GTP-binding; Myristate.
SQ SEQUENCE 206 AA; 23454 MW; 65D5DDE4FB748072 CRC64;

Query Match          48.98; Score 1108; DB 15; Length 206;
Best Local Similarity 97.18; Pred. No. 1.8e-81;
Matches 200; Conservative 3; Mismatches 3; Indels 0; Gaps 0;

Qy 112 MGGKWSKSSVVGWPTVTRRMRRAEPAADGGAASRDLEKKGGLHSGRQDILDLWIYHTQY 171
Db 1 MGGKWSKSSVVGWPAVRERMRRAEPAADGGAASRDLEKKGGLHSGRQDILDLWIYHTQY 60

Qy 172 QEEVEGFPVTPQVPLRPMYKAAVDLSHFLKEKGGLEGLHSGRQDILDLWIYHTQY 231
Db 61 QEEVEGFPVTPQVPLRPMYKAAVDLSHFLKEKGGLEGLHSGRQDILDLWIYHTQY 120

Qy 232 FPDQNYTPGCGVRYPLTFGWCYKLVPEPKVEANKGENTSLHPVSLHGMDDPEREV 291
Db 121 FPDQNYTPGCGVRYPLTFGWCYKLVPEPKVEANKGENTSLHPVSLHGMDDPEREV 180

Qy 292 LEWRFDSRLAFHHVARELHPEYFKNC 317
Db 181 LEWRFDSRLAFHHVARELHPEYFKNC 206

RESULT 6
Q9WM16 PRELIMINARY; PRT; 206 AA.
AC Q9WM16;
DT 01-NOV-1999 (TrEMBLrel. 12, Created)
DT 01-NOV-1999 (TrEMBLrel. 12, Last sequence update)
DT 01-DEC-2001 (TrEMBLrel. 19, Last annotation update)
DE NEGATIVE FACTOR (F-PROTEIN) (27 KDA PROTEIN).
GN NEF.
OS Human immunodeficiency virus type 1.
OC Viruses; Retroid viruses; Retroviridae; Lentivirus.
OX NCBI_TaxID=11676;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=7-IT;
RA Visco Comandini U., Yun Z., Paganelli R., Orlandi P., Salotti A.,
RA Johansson B., Vahne A., Sonnerborg A.;
RT "HIV-1 nef mutations and clinical long-term non progression: a
RT molecular epidemiology study.";
RL Submitted (JUN-1997) to the EMBL/GenBank/DBJ databases.
CC -1- SIMILARITY: NEF HAS GTPASE, GTP-BINDING AND AUTOPHOSPHORYLATING
CC ACTIVITIES, IT SEEMS TO DOWN-REGULATE THE CD4(T4) ANTIGEN.
DR EMBL; AF011494; AAD01472.1; -.
DR HSP; P03406; IEFN.
DR InterPro: IPR001558; F-protein.
DR Pfam: PF00469; F-protein; 1.
DR AIDS; GTP-binding; Myristate.
KW AIDS; GTP-binding; Myristate.
SQ SEQUENCE 206 AA; 23333 MW; 2EBF2A6A3ECAF5EA CRC64;

Query Match          48.88; Score 1105; DB 15; Length 206;
Best Local Similarity 97.18; Pred. No. 3.1e-81;
Matches 200; Conservative 2; Mismatches 4; Indels 0; Gaps 0;

Qy 112 MGGKWSKSSVVGWPTVTRRMRRAEPAADGGAASRDLEKKGGLHSGRQDILDLWIYHTQY 171
Db 1 MGGKWSKSSVVGWPAVRERMRRAEPAADGGAASRDLEKKGGLHSGRQDILDLWIYHTQY 60

Qy 172 QEEVEGFPVTPQVPLRPMYKAAVDLSHFLKEKGGLEGLHSGRQDILDLWIYHTQY 231
Db 61 QEEVEGFPVTPQVPLRPMYKAAVDLSHFLKEKGGLEGLHSGRQDILDLWIYHTQY 120

Qy 232 FPDQNYTPGCGVRYPLTFGWCYKLVPEPKVEANKGENTSLHPVSLHGMDDPEREV 291
Db 121 FPDQNYTPGCGVRYPLTFGWCYKLVPEPKVEANKGENTSLHPVSLHGMDDPEREV 180

Qy 292 LEWRFDSRLAFHHVARELHPEYFKNC 317
Db 181 LEWRFDSRLAFHHVARELHPEYFKNC 206

RESULT 7
Q9WM30 PRELIMINARY; PRT; 206 AA.
AC Q9WM30;
DT 01-NOV-1999 (TrEMBLrel. 12, Created)
DT 01-NOV-1999 (TrEMBLrel. 12, Last sequence update)
DT 01-DEC-2001 (TrEMBLrel. 19, Last annotation update)
DE NEGATIVE FACTOR (F-PROTEIN) (27 KDA PROTEIN).
GN NEF.
OS Human immunodeficiency virus type 1.
OC Viruses; Retroid viruses; Retroviridae; Lentivirus.
OX NCBI_TaxID=11676;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=8-IT;
RA Visco Comandini U., Yun Z., Paganelli R., Orlandi P., Salotti A.,
RA Johansson B., Vahne A., Sonnerborg A.;
RT "HIV-1 nef mutations and clinical long-term non progression: a
RT molecular epidemiology study.";
RL Submitted (JUN-1997) to the EMBL/GenBank/DBJ databases.
CC -1- SIMILARITY: NEF HAS GTPASE, GTP-BINDING AND AUTOPHOSPHORYLATING
CC ACTIVITIES, IT SEEMS TO DOWN-REGULATE THE CD4(T4) ANTIGEN.
```

```
DR EMBL; AF011469; AAD01447.1; -
DR HSP; P03406; IEFN.
DR InterPro; IPR001558; F-protein.
DR Pfam; PF00469; F-protein; 1.
DR AIDS; GTP-binding; Myristate.
SQ SEQUENCE 206 AA; 23345 MW; 0ED69927C2E03BB6 CRC64;

Query Match 48.8%; Score 1104; DB 15; Length 206;
Best Local Similarity 97.1%; Pred. No. 3.7e-81;
Matches 200; Conservative 1; Mismatches 5; Indels 0; Gaps 0;

Qy 112 MGGKWSKSSVVGWPTVRMRRAEPAADGVGAASRDLEKHAITSNTAATNAACAWLEA 171
Db 1 MGGKWSKSSVVGWPAVRMRRAEPAADGVGAASRDLEKHAITSNTAANNAACAWLEA 60

Qy 172 QEEVEGFPVTPQVPLRPMTYKAAVDLSHFLKEKGGLEGLHSQRRODILDWIYHTQGY 231
Db 61 QEEVEGFPVTPQVPLRPMTYKAAVDLSHFLKEKGGLEGLHSQRRODILDWIYHTQGY 120

Qy 232 FPDQNTYTPGCGVRYPLTFGWCYKLVPEPKVEEANKGENTSLHHPVSLHGMDDPEREV 291
Db 121 FPDQNTYTPGCGVRYPLTFGWCYKLVPEPKVEEANKGENTSLHHPVSLHGMDDPEREV 180

Qy 292 LEWRFDLSRLAFHHVARELHPEYFKNC 317
Db 181 LEWRFDLSRLAFHHVARELHPEYFKNC 206

RESULT 8
Q9PXW8 PRELIMINARY; PRT; 206 AA.
AC Q9PXW8
DT 01-MAY-2000 (TrEMBLrel. 13, Created)
DT 01-MAY-2000 (TrEMBLrel. 13, Last sequence update)
DT 01-DEC-2001 (TrEMBLrel. 19, Last annotation update)
DE NEGATIVE FACTOR (F-PROTEIN) (27 KDA PROTEIN).
GN NEF.
OS Human immunodeficiency virus type 1.
OC Viruses; Retroid viruses; Retroviridae; Lentivirus.
OX NCBI_TaxID=11676;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=93152025; PubMed=1301062;
RA Kohleisen B., Neumann M., Herrmann R., Brack-Werner R., Krohn K.J.,
RA Ovod V., Ranki A., Erfle V.;
RT "Cellular localization of Nef expressed in persistently HIV-1-infected
RT low-producer astrocytes.";
RL AIDS 6:1427-1436(1992).
CC -1- SIMILARITY: NEF HAS GTPASE, GTP-BINDING AND AUTOPHOSPHORYLATING
CC HSP; P03406; IEFN.
DR InterPro; IPR001558; F-protein.
DR Pfam; PF00469; F-protein; 1.
KW AIDS; GTP-binding; Myristate.
SQ SEQUENCE 206 AA; 23439 MW; 8478CF194D39D5B9 CRC64;

Query Match 48.6%; Score 1101; DB 15; Length 206;
Best Local Similarity 96.1%; Pred. No. 6.5e-81;
Matches 198; Conservative 4; Mismatches 4; Indels 0; Gaps 0;

Qy 112 MGGKWSKSSVVGWPTVRMRRAEPAADGVGAASRDLEKHAITSNTAATNAACAWLEA 171
Db 1 MGGKWSKSSVVGWPAVRMRRAEPAADGVGAASRDLEKHAITSNTAANNAACAWLEA 60

Qy 172 QEEVEGFPVTPQVPLRPMTYKAAVDLSHFLKEKGGLEGLHSQRRODILDWIYHTQGY 231
Db 61 QEEVEGFPVTPQVPLRPMTYKAAVDLSHFLKEKGGLEGLHSQRRODILDWIYHTQGY 120

Qy 232 FPDQNTYTPGCGVRYPLTFGWCYKLVPEPKVEEANKGENTSLHHPVSLHGMDDPEREV 291
Db 121 FPDQNTYTPGCGVRYPLTFGWCYKLVPEPKVEEANKGENTSLHHPVSLHGMDDPEREV 180
```

```
Qy 292 LEWRFDLSRLAFHHVARELHPEYFKNC 317
Db 181 LEWRFDLSRLAFHHVARELHPEYFKNC 206

RESULT 9
Q90179 PRELIMINARY; PRT; 206 AA.
AC Q90179
DT 01-NOV-1998 (TrEMBLrel. 08, Created)
DT 01-NOV-1998 (TrEMBLrel. 08, Last sequence update)
DT 01-DEC-2001 (TrEMBLrel. 19, Last annotation update)
DE NEGATIVE FACTOR (F-PROTEIN) (27 KDA PROTEIN).
GN NEF.
OS Human immunodeficiency virus type 1.
OC Viruses; Retroid viruses; Retroviridae; Lentivirus.
OX NCBI_TaxID=11676;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=95074930; PubMed=7983770;
RA Fang H., Pincus S.H.;
RT "Unique insertion sequence and pattern of CD4 expression in variants
RT selected with immunotoxins from human immunodeficiency virus type 1-
RT infected T cells.";
RL J. Virol. 69:75-81(1995).
RN [2]
RP SEQUENCE FROM N.A.
RA Fang H., Pincus S.H.;
RT "Spontaneous activation of human immunodeficiency virus type 1 in an
RT immunotoxin-resistant variant T cell line.";
RL Submitted (JUN-1998) to the EMBL/GenBank/DBJ databases.
CC -1- SIMILARITY: NEF HAS GTPASE, GTP-BINDING AND AUTOPHOSPHORYLATING
CC ACTIVITIES, IT SEEMS TO DOWN-REGULATE THE CD4(T4) ANTIGEN.
DR EMBL; AF070521; AAC28453.1;
DR HSP; P03406; IEFN.
DR InterPro; IPR001558; F-protein.
DR Pfam; PF00469; F-protein; 1.
KW AIDS; GTP-binding; Myristate.
SQ SEQUENCE 206 AA; 23334 MW; 3E2B9C4017FDC68 CRC64;

Query Match 48.4%; Score 1096; DB 15; Length 206;
Best Local Similarity 95.6%; Pred. No. 1.6e-80;
Matches 197; Conservative 4; Mismatches 5; Indels 0; Gaps 0;

Qy 112 MGGKWSKSSVVGWPTVRMRRAEPAADGVGAASRDLEKHAITSNTAATNAACAWLEA 171
Db 1 MGGKWSKSSVVGWPAVRMRRAEPAADGVGAASRDLEKHAITSNTAANNAACAWLEA 60

Qy 172 QEEVEGFPVTPQVPLRPMTYKAAVDLSHFLKEKGGLEGLHSQRRODILDWIYHTQGY 231
Db 61 QEEVEGFPVTPQVPLRPMTYKAAVDLSHFLKEKGGLEGLHSQRRODILDWIYHTQGY 120

Qy 232 FPDQNTYTPGCGVRYPLTFGWCYKLVPEPKVEEANKGENTSLHHPVSLHGMDDPEREV 291
Db 121 FPDQNTYTPGCGVRYPLTFGWCYKLVPEPKVEEANKGENTSLHHPVSLHGMDDPEREV 180

Qy 292 LEWRFDLSRLAFHHVARELHPEYFKNC 317
Db 181 LEWRFDLSRLAFHHVARELHPEYFKNC 206

RESULT 10
Q85588 PRELIMINARY; PRT; 206 AA.
AC Q85588
DT 01-NOV-1996 (TrEMBLrel. 01, Created)
DT 01-NOV-1996 (TrEMBLrel. 01, Last sequence update)
DT 01-DEC-2001 (TrEMBLrel. 19, Last annotation update)
DE NEGATIVE FACTOR (F-PROTEIN) (27 KDA PROTEIN).
GN NEF.
OS Aids-associated retrovirus.
```

Viruses; Retroviral viruses; Retroviridae.

OC NCBI\_TaxID=11966;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RX MEDLINE=86067228; PubMed=2999715;  
 RA Ratner L., Starich B., Josephs S.F., Hahn B.H., Reddy E.P., Arya S.K.,  
 RA Livak K.J., Petteway S.R.Jr., Pearson M.L., Haseltine W.A., Arya S.K.,  
 RA Wong-staal F.;  
 RT "Polymorphism of the 3' open reading frame of the virus associated  
 RT with the acquired immune deficiency syndrome, human T-lymphotropic  
 RT virus type III";  
 RL Nucleic Acids Res. 13:8219-8229(1985).  
 CC -1- SIMILARITY: NEF HAS GTPASE, GTP-BINDING AND AUTOPHOSPHORYLATING  
 CC ACTIVITIES. IT SEEMS TO DOWN-REGULATE THE CD4(T4) ANTIGEN.  
 DR EMBL; X03188; CAA26947.1; -;  
 DR HSP; P03406; 1EFN.  
 DR InterPro; IPR001558; F-protein.  
 DR Pfam; PF00469; F-protein; 1.  
 DR AIDS; GTP-binding; Myristate.  
 KW AIDS; GTP-binding; Myristate.  
 SQ SEQUENCE 206 AA; 23419 MW; 3B25EB332A479A46 CRC64;  
 Query Match 48.4%; Score 1096; DB 15; Length 206;  
 Best Local Similarity 96.6%; Pred. No. 1.6e-80;  
 Matches 199; Conservative 4; Mismatches 3; Indels 0; Gaps 0;  
 QY 112 MGGKWSKSSVVGWPTVRRMRRAEPAADGVGAASRDLEKKGAISSNTAATNAACAWLEA 171  
 Db 1 MGGKWSKSSVVGWPAVRMRRAEPAADGVGAASRDLEKKGAISSNTAATNAACAWLEA 60  
 QY 172 QEEVEGFPVTPQVPLRPMTYKAAVDLSHFLKKEGGLIHSORRODILDWIYHTQY 231  
 Db 61 QEEVEGFPVTPQVPLRPMTYKAAVDLSHFLKKEGGLIHSORRODILDWIYHTQY 120  
 QY 232 FPDQNTPTGPGVRYPLTFGWCYKLVPEPKVEANKGENTSLHPVSLHGMDDPREV 291  
 Db 121 FPDQNTPTGPGVRYPLTFGWCYKLVPEPKVEANKGENTSLHPVSLHGMDDPREV 180  
 QY 292 LEWRDLSLAFHHVARELHPEYFKNK 317  
 Db 181 LEWRDLSLAFHHVARELHPEYFKNK 206  
 RESULT 11  
 Q78244  
 ID Q78244 PRELIMINARY; PRT; 206 AA.  
 AC Q78244;  
 DT 01-NOV-1996 (TrEMBLrel. 01, Created)  
 DT 01-NOV-1996 (TrEMBLrel. 01, Last sequence update)  
 DT 01-DEC-2001 (TrEMBLrel. 19, Last annotation update)  
 DE NEGATIVE FACTOR (F-PROTEIN) (27 KDA PROTEIN).  
 GN NEF.  
 OS Human immunodeficiency virus type 1.  
 OC Viruses; Retroviral viruses; Retroviridae; Lentivirus.  
 OX NCBI\_TaxID=11676;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RA Carlini F., Federico M., Equestre M., Ricci S., Ratti G., Zibai Q.,  
 RA Verani P., Rossi G.B.;  
 RT "Sequence analysis of HIV-1 proviral DNA from a non producer  
 RT chronically infected HUT-78 cellular clone";  
 RL J. Viral Diseases 1:40-55(1992).  
 RN [2]  
 RP SEQUENCE FROM N.A.  
 RX MEDLINE=89352106; PubMed=2765297;  
 RA Federico M., Titti F., Butto S., Orecchia A., Carlini F., Taddeo B.,  
 RA Macchi B., Mangiano N., Verani P., Rossi G.;  
 RT "Biological and molecular characterization of producer and non  
 RT producer clones from HUT-78 infected with a patient HIV isolate";  
 RL AIDS Res. Hum. Retroviruses 5:385-396(1989).  
 RN [3]  
 RP SEQUENCE FROM N.A.  
 RA Titti F., Federico M., Butto S., Orecchia A., Carlini F., Taddeo B.,

RA Borsetti A., Saggio I., Verani P., Rossi G.;  
 RT "Variability of HIV-1 virus: Characteristics of an infected but not  
 RT productive clone";  
 RL Int. J. Immunopharmacol. 3:17-23(1990).  
 CC -1- SIMILARITY: NEF HAS GTPASE, GTP-BINDING AND AUTOPHOSPHORYLATING  
 CC ACTIVITIES. IT SEEMS TO DOWN-REGULATE THE CD4(T4) ANTIGEN.  
 DR EMBL; Z11530; CAA77629.1; -;  
 DR HSP; P03406; 1EFN.  
 DR InterPro; IPR001558; F-protein.  
 DR Pfam; PF00469; F-protein; 1.  
 DR AIDS; GTP-binding; Myristate.  
 KW AIDS; GTP-binding; Myristate.  
 SQ SEQUENCE 206 AA; 23398 MW; A31CC7244BBB2B43 CRC64;

Query Match 48.4%; Score 1096; DB 15; Length 206;  
 Best Local Similarity 96.6%; Pred. No. 1.6e-80;  
 Matches 199; Conservative 3; Mismatches 4; Indels 0; Gaps 0;

QY 112 MGGKWSKSSVVGWPTVRRMRRAEPAADGVGAASRDLEKKGAISSNTAATNAACAWLEA 171  
 Db 1 MGGKWSKSSVVGWPAVRMRRAEPAADGVGAASRDLEKKGAISSNTAATNAACAWLEA 60  
 QY 172 QEEVEGFPVTPQVPLRPMTYKAAVDLSHFLKKEGGLIHSORRODILDWIYHTQY 231  
 Db 61 QEEVEGFPVTPQVPLRPMTYKAAVDLSHFLKKEGGLIHSORRODILDWIYHTQY 120  
 QY 232 FPDQNTPTGPGVRYPLTFGWCYKLVPEPKVEANKGENTSLHPVSLHGMDDPREV 291  
 Db 121 FPDQNTPTGPGVRYPLTFGWCYKLVPEPKVEANKGENTSLHPVSLHGMDDPREV 180  
 QY 292 LEWRDLSLAFHHVARELHPEYFKNK 317  
 Db 181 LEWRDLSLAFHHVARELHPEYFKNK 206

RESULT 12  
 Q9WLM7  
 ID Q9WLM7 PRELIMINARY; PRT; 206 AA.  
 AC Q9WLM7;  
 DT 01-NOV-1999 (TrEMBLrel. 12, Created)  
 DT 01-NOV-1999 (TrEMBLrel. 12, Last sequence update)  
 DT 01-DEC-2001 (TrEMBLrel. 19, Last annotation update)  
 DE NEGATIVE FACTOR (F-PROTEIN) (27 KDA PROTEIN).  
 GN NEF.  
 OS Human immunodeficiency virus type 1.  
 OC Viruses; Retroviral viruses; Retroviridae; Lentivirus.  
 OX NCBI\_TaxID=11676;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RA Visco Comandini U., Yun Z., Paganelli R., Orlandi P., Salotti A.,  
 RA Johansson B., Vahine A., Sonnerborg A.;  
 RT "HIV-1 nef mutations and clinical long-term non progression: a  
 RT molecular epidemiology study";  
 RL Submitted (FEB-1998) to the EMBL/GenBank/DBJ databases.  
 CC -1- SIMILARITY: NEF HAS GTPASE, GTP-BINDING AND AUTOPHOSPHORYLATING  
 CC ACTIVITIES. IT SEEMS TO DOWN-REGULATE THE CD4(T4) ANTIGEN.  
 DR EMBL; AF047084; AAD02458.1; -;  
 DR HSP; P03406; 1EFN.  
 DR InterPro; IPR001558; F-protein.  
 DR Pfam; PF00469; F-protein; 1.  
 DR AIDS; GTP-binding; Myristate.  
 KW AIDS; GTP-binding; Myristate.  
 SQ SEQUENCE 206 AA; 23512 MW; E09E3BEF828A83C0 CRC64;

Query Match 48.2%; Score 1092; DB 15; Length 206;  
 Best Local Similarity 95.6%; Pred. No. 3.5e-80;  
 Matches 197; Conservative 7; Mismatches 2; Indels 0; Gaps 0;

QY 112 MGGKWSKSSVVGWPTVRRMRRAEPAADGVGAASRDLEKKGAISSNTAATNAACAWLEA 171  
 Db 1 MGGKWSKSSVVGWPTVRRMRRAEPAADGVGAASRDLEKKGAISSNTAATNAACAWLEA 60

```
OS 172 QEEVEGFPVTPQVPLRPMYKAAVDLSHFLKEKGLGSLHSQRQDILDLDWIYHTQGY 231
OC ||||||| ||||||| ||||||| ||||||| ||||||| ||||||| ||||||| |||||||
OX 61 QEEVEGFPVTPQVPLRPMYKAAVDLSHFLKEKGLGSLHSQRQDILDLDWIYHTQGY 120
RN ||||||| ||||||| ||||||| ||||||| ||||||| ||||||| ||||||| |||||||
RA 232 FPDWQNYTPGCVRYPLTFGWCYKLVPEPKVEEANKGENTSLHLPVSLHGMDDPEREV 291
RP ||||||| ||||||| ||||||| ||||||| ||||||| ||||||| ||||||| |||||||
DB 121 FPDWQNYTPGCVRYPLTFGWCYKLVPEPKVEEANKGENTSLHLPVSLHGMDDPEREV 180
QY 292 LEWRFDSRLAFHHVARELHPEYFKNC 317
DB ||||||| ||||||| ||||||| ||||||| ||||||| ||||||| ||||||| |||||||
DB 181 LEWRFDSRLAFHHVARELHPEYFKNC 206

RESULT 13
Q74905 PRELIMINARY; PRT; 206 AA.
ID Q74905;
AC Q74905;
DT 01-NOV-1996 (TReMBLrel. 01, Created)
DT 01-NOV-1996 (TReMBLrel. 01, Last sequence update)
DT 01-DEC-2001 (TReMBLrel. 19, Last annotation update)
DE NEGATIVE FACTOR (F-PROTEIN) (27 KDA PROTEIN).
GN NEF.
OS Human immunodeficiency virus type 1.
OC Viruses; Retroid viruses; Retroviridae; Lentivirus.
OX NCBI_TaxID=11676;
RN ||||||| ||||||| ||||||| ||||||| ||||||| ||||||| ||||||| |||||||
RA 112 MGGKWSKSVVGVPTVRMRRAEPAADGVGAASRDLEKHGAITSNTAATNAACAWLEA 171
RP ||||||| ||||||| ||||||| ||||||| ||||||| ||||||| ||||||| |||||||
DB 1 MGGKWSKSVVGVPTVRMRRAEPAADGVGAASRDLEKHGAITSNTAATNAACAWLEA 60

Query Match 48.1%; Score 1090; DB 15; Length 206;
Best Local Similarity 95.1%; Pred. No. 5e-80; Indels 0; Gaps 0;
Matches 196; Conservative 6; Mismatches 4;

OS 172 QEEVEGFPVTPQVPLRPMYKAAVDLSHFLKEKGLGSLHSQRQDILDLDWIYHTQGY 231
OC ||||||| ||||||| ||||||| ||||||| ||||||| ||||||| ||||||| |||||||
OX 61 QEEVEGFPVTPQVPLRPMYKAAVDLSHFLKEKGLGSLHSQRQDILDLDWIYHTQGY 120
RN ||||||| ||||||| ||||||| ||||||| ||||||| ||||||| ||||||| |||||||
RA 232 FPDWQNYTPGCVRYPLTFGWCYKLVPEPKVEEANKGENTSLHLPVSLHGMDDPEREV 291
RP ||||||| ||||||| ||||||| ||||||| ||||||| ||||||| ||||||| |||||||
DB 121 FPDWQNYTPGCVRYPLTFGWCYKLVPEPKVEEANKGENTSLHLPVSLHGMDDPEREV 180
QY 292 LEWRFDSRLAFHHVARELHPEYFKNC 317
DB ||||||| ||||||| ||||||| ||||||| ||||||| ||||||| ||||||| |||||||
DB 181 LEWRFDSRLAFHHVARELHPEYFKNC 206

RESULT 14
Q89561 PRELIMINARY; PRT; 206 AA.
ID Q89561;
AC Q89561;
DT 01-NOV-1996 (TReMBLrel. 01, Created)
DT 01-NOV-1996 (TReMBLrel. 01, Last sequence update)
DT 01-DEC-2001 (TReMBLrel. 19, Last annotation update)
DE NEGATIVE FACTOR (F-PROTEIN) (27 KDA PROTEIN).
GN NEF.
```

```
OS Human immunodeficiency virus type 1.
OC Viruses; Retroid viruses; Retroviridae; Lentivirus.
OX NCBI_TaxID=11676;
RN ||||||| ||||||| ||||||| ||||||| ||||||| ||||||| ||||||| |||||||
RA Arens M.O., Ratner L., Joseph T., Bandres J.;
RP Submitted (JAN-1996) to the EMBL/GenBank/DBJ databases.
DB -1- SIMILARITY: NEF HAS GTPASE, GTP-BINDING AND AUTOPHOSPHORYLATING
CC ACTIVITIES, IT SEEMS TO DOWN-REGULATE THE CD4(T4) ANTIGEN.
CC EMBL; U44448; AAB38200.1; -.
DR EMBL; U44447; AAB38199.1; -.
DR HSP; Q70627; 2NEF.
DR InterPro; IPR001558; F-protein.
DR Pfam; PF00469; F-protein; 1.
DR AIDS; GTP-binding; Myristate.
KW AIDS; GTP-binding; Myristate.
SQ SEQUENCE 206 AA; 23629 MW; BF467FB1B7147CDE CRC64;

Query Match 48.0%; Score 1087; DB 15; Length 206;
Best Local Similarity 94.7%; Pred. No. 8.7e-80; Indels 0; Gaps 0;
Matches 195; Conservative 7; Mismatches 4;

OS 112 MGGKWSKSVVGVPTVRMRRAEPAADGVGAASRDLEKHGAITSNTAATNAACAWLEA 171
OC ||||||| ||||||| ||||||| ||||||| ||||||| ||||||| ||||||| |||||||
DB 1 MGGKWSKSVVGVPTVRMRRAEPAADGVGAASRDLEKHGAITSNTAATNAACAWLEA 60

OS 172 QEEVEGFPVTPQVPLRPMYKAAVDLSHFLKEKGLGSLHSQRQDILDLDWIYHTQGY 231
OC ||||||| ||||||| ||||||| ||||||| ||||||| ||||||| ||||||| |||||||
DB 61 QEEVEGFPVTPQVPLRPMYKAAVDLSHFLKEKGLGSLHSQRQDILDLDWIYHTQGY 120
QY 232 FPDWQNYTPGCVRYPLTFGWCYKLVPEPKVEEANKGENTSLHLPVSLHGMDDPEREV 291
DB ||||||| ||||||| ||||||| ||||||| ||||||| ||||||| ||||||| |||||||
DB 121 FPDWQNYTPGCVRYPLTFGWCYKLVPEPKVEEANKGENTSLHLPVSLHGMDDPEREV 180
QY 292 LEWRFDSRLAFHHVARELHPEYFKNC 317
DB ||||||| ||||||| ||||||| ||||||| ||||||| ||||||| ||||||| |||||||
DB 181 LEWRFDSRLAFHHVARELHPEYFKNC 206

RESULT 15
Q74913 PRELIMINARY; PRT; 206 AA.
ID Q74913;
AC Q74913;
DT 01-NOV-1996 (TReMBLrel. 01, Created)
DT 01-NOV-1996 (TReMBLrel. 01, Last sequence update)
DT 01-DEC-2001 (TReMBLrel. 19, Last annotation update)
DE NEGATIVE FACTOR (F-PROTEIN) (27 KDA PROTEIN).
GN NEF.
OS Human immunodeficiency virus type 1.
OC Viruses; Retroid viruses; Retroviridae; Lentivirus.
OX NCBI_TaxID=11676;
RN ||||||| ||||||| ||||||| ||||||| ||||||| ||||||| ||||||| |||||||
RA Ratner L., Joseph T., Bandres J., Ghosh S., Heyden N.V., Templeton A.,
RP Hahn B., Powderly W., Arens M.;
DB "Sequence heterogeneity of Nef transcripts in HIV-1-infected subjects
CC at different stages of disease.";
CC Virology 223:245-250(1996).
CC -1- SIMILARITY: NEF HAS GTPASE, GTP-BINDING AND AUTOPHOSPHORYLATING
CC ACTIVITIES, IT SEEMS TO DOWN-REGULATE THE CD4(T4) ANTIGEN.
CC EMBL; U44453; AAB38205.1; -.
DR HSP; P03406; IEFN.
DR InterPro; IPR001558; F-protein.
DR Pfam; PF00469; F-protein; 1.
DR AIDS; GTP-binding; Myristate.
KW AIDS; GTP-binding; Myristate.
SQ SEQUENCE 206 AA; 23649 MW; E5B997FA0B70FC175 CRC64;

Query Match 47.9%; Score 1085; DB 15; Length 206;
Best Local Similarity 95.6%; Pred. No. 1.3e-79;
Matches 196; Conservative 5; Mismatches 4; Indels 0; Gaps 0;
```





SQ SEQUENCE 206 AA; 23648 MW; E5BD6FA0B70FC175 CRC64;

Query Match 47.7%; Score 1079; DB 15; Length 206;  
Best Local Similarity 95.6%; Pred. No. 3.8e-79;  
Matches 195; Conservative 5; Mismatches 4; Indels 0; Gaps 0;

QY 112 MGGKSKSSVGVWPTVRMRRAEPAADGVGAASRDLEKHGAITSNTAATNAACAWLEA 171  
|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||  
Db 1 MGGKSKSSVGVWPTVRMRRAEPAADGVGAASRDLEKHGAITSNTAATNAACAWLEA 60  
QY 172 QEEVEGVFPTQVPLRPMTYKAAVDLSHFLKEKGLEGLIHSQRQDILDWIYHTQY 231  
|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||  
Db 61 QEEVEGVFPTQVPLRPMTYKAAVDLSHFLKEKGLEGLIHSQRQDILDWIYHTQY 120  
QY 232 FPDWQNTPGVRYPLTFGWCYKLVPEPKVEEANKGENTSLHPVSLHGMDDPEREV 291  
|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||  
Db 121 FPDWQNTPGVRYPLTFGWCYKLVPEPKVEEANKGENTSLHPVSLHGMDDPEREV 180  
QY 292 LEWRFSRLAFHHVARELHPEYFK 315  
|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||  
Db 181 LEWRFSRLAFHHVARELHPEYFK 204

RESULT 19  
Q74909 PRELIMINARY; PRT: 206 AA.  
AC Q74909;  
DT 01-NOV-1996 (Tremblrel. 01, Created)  
DT 01-NOV-1996 (Tremblrel. 01, Last sequence update)  
DT 01-DEC-2001 (Tremblrel. 19, Last annotation update)  
DE NEGATIVE FACTOR (F-PROTEIN) (27 KDA PROTEIN).  
GN NEF.  
OS Human immunodeficiency virus type 1.  
OC Viruses; Retroviral viruses; Retroviridae; Lentivirus.  
OX NCBI\_TaxID=11676;  
RN [1]  
RX MEDLINE=96400183; PubMed=8806559;  
RA Rathel L., Joseph T., Bandres J., Ghosh S., Heyden N.V., Templeton A.,  
Hahn B., Powderly W., Arens M.;  
RT "Sequence heterogeneity of Nef transcripts in HIV-1-infected subjects  
at different stages of disease";  
RL Virology 223:245-250(1996).  
CC -!- SIMILARITY: NEF HAS GTPASE, GTP-BINDING AND AUTOPHOSPHORYLATING  
ACTIVITIES, IT SEEMS TO DOWN-REGULATE THE CD4(T4) ANTIGEN.  
DR EMBL: U44449; AAB38201.1; -;  
DR HSP: Q70627; 2NEF.  
DR InterPro: IPR001558; F-protein.  
DR Pfam: PF00469; F-protein; 1.  
KW AIDS; GTP-binding; Myristate.  
SQ SEQUENCE 206 AA; 23488 MW; F0596D6FAAA81A05 CRC64;

Query Match 47.6%; Score 1077; DB 15; Length 206;  
Best Local Similarity 94.2%; Pred. No. 5.6e-79;  
Matches 194; Conservative 7; Mismatches 5; Indels 0; Gaps 0;

QY 112 MGGKSKSSVGVWPTVRMRRAEPAADGVGAASRDLEKHGAITSNTAATNAACAWLEA 171  
|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||  
Db 1 MGGKSKSSVGVWPTVRMRRAEPAADGVGAASRDLEKHGAITSNTAATNAACAWLEA 60  
QY 172 QEEVEGVFPTQVPLRPMTYKAAVDLSHFLKEKGLEGLIHSQRQDILDWIYHTQY 231  
|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||  
Db 61 QEEVEGVFPTQVPLRPMTYKAAVDLSHFLKEKGLEGLIHSQRQDILDWIYHTQY 120  
QY 232 FPDWQNTPGVRYPLTFGWCYKLVPEPKVEEANKGENTSLHPVSLHGMDDPEREV 291  
|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||  
Db 121 FPDWQNTPGVRYPLTFGWCYKLVPEPKVEEANKGENTSLHPVSLHGMDDPEREV 180  
QY 292 LEWRFSRLAFHHVARELHPEYFKNC 317  
|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||  
Db 181 LEWRFSRLAFHHVARELHPEYFKNC 206

RESULT 20  
Q9QPN3 PRELIMINARY; PRT: 202 AA.  
AC Q9QPN3;  
DT 01-MAY-2000 (Tremblrel. 13, Created)  
DT 01-MAY-2000 (Tremblrel. 13, Last sequence update)  
DT 01-DEC-2001 (Tremblrel. 19, Last annotation update)  
DE NEGATIVE FACTOR (F-PROTEIN) (27 KDA PROTEIN) (FRAGMENT).  
GN NEF.  
OS Human immunodeficiency virus type 1.  
OC Viruses; Retroviral viruses; Retroviridae; Lentivirus.  
OX NCBI\_TaxID=11676;  
RN [1]  
RP SEQUENCE FROM N.A.  
RC STRAIN=HIV-11AI;  
RA Piedade J., Esteves A., Parreira R., Venenno T., Barros M.F.,  
Canas-Ferreira W.F.;  
RT "Cloning and expression of HIV-1 nef gene in the carrier-adjutant  
pVUB3 expression system based on the major lipoprotein (Opr1) from the  
outer membrane of Pseudomonas aeruginosa";  
RL Submitted (JUL-1999) to the EMBL/GenBank/DBJ databases.  
CC -!- SIMILARITY: NEF HAS GTPASE, GTP-BINDING AND AUTOPHOSPHORYLATING  
ACTIVITIES, IT SEEMS TO DOWN-REGULATE THE CD4(T4) ANTIGEN.  
DR EMBL: AF166101; AAD47831.1; -;  
DR HSP: P03406; IEFN.  
DR InterPro: IPR001558; F-protein.  
DR Pfam: PF00469; F-protein; 1.  
KW AIDS; GTP-binding; Myristate.  
FT NON\_TER 202  
SQ SEQUENCE 202 AA; 23033 MW; CD61DFA6F386CC89 CRC64;

Query Match 47.5%; Score 1076; DB 15; Length 202;  
Best Local Similarity 97.5%; Pred. No. 6.5e-79;  
Matches 197; Conservative 0; Mismatches 5; Indels 0; Gaps 0;

QY 112 MGGKSKSSVGVWPTVRMRRAEPAADGVGAASRDLEKHGAITSNTAATNAACAWLEA 171  
|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||  
Db 1 MGGKSKSSVGVWPTVRMRRAEPAADGVGAASRDLEKHGAITSNTAATNAACAWLEA 60  
QY 172 QEEVEGVFPTQVPLRPMTYKAAVDLSHFLKEKGLEGLIHSQRQDILDWIYHTQY 231  
|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||  
Db 61 QEEVEGVFPTQVPLRPMTYKAAVDLSHFLKEKGLEGLIHSQRQDILDWIYHTQY 120  
QY 232 FPDWQNTPGVRYPLTFGWCYKLVPEPKVEEANKGENTSLHPVSLHGMDDPEREV 291  
|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||  
Db 121 FPDWQNTPGVRYPLTFGWCYKLVPEPKVEEANKGENTSLHPVSLHGMDDPEREV 180  
QY 292 LEWRFSRLAFHHVARELHPEY 313  
|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||  
Db 181 LEWRFSRLAFHHVARELHPEY 202

RESULT 21  
Q9Q596 PRELIMINARY; PRT: 206 AA.  
AC Q9Q596;  
DT 01-MAY-2000 (Tremblrel. 13, Created)  
DT 01-MAY-2000 (Tremblrel. 13, Last sequence update)  
DT 01-DEC-2001 (Tremblrel. 19, Last annotation update)  
DE NEGATIVE FACTOR (F-PROTEIN) (27 KDA PROTEIN).  
GN NEF.  
OS Human immunodeficiency virus type 1.  
OC Viruses; Retroviral viruses; Retroviridae; Lentivirus.  
OX NCBI\_TaxID=11676;  
RN [1]  
RP SEQUENCE FROM N.A.  
RA Wang B., Saksena N.K.;  
RT "HIV-1 Strains from a cohort of American subjects reveal the presence  
of a V2 region extension unique to slow progressors and non-  
progressors";

```

RL AIDS 0.0-0(2000).
CC -1- SIMILARITY: NEF HAS GTPASE, GTP-BINDING AND AUTOPHOSPHORYLATING
CC ACTIVITIES, IT SEEMS TO DOWN-REGULATE THE CD4(T4) ANTIGEN.
DR EMBL: AF203197; AAF25319.1; -
DR HSP: P03406; IEFN.
DR InterPro: IPR001558; F-protein.
DR Pfam: PF00469; F-protein; 1.
KW AIDS; Gtp-binding; Myristate.
SQ SEQUENCE 206 AA; 23510 MW; 081DF3A12ESA7576 CRC64;

Query Match 47.3%; Score 1072; DB 15; Length 206;
Best Local Similarity 94.7%; Pred. No. 1.4e-78;
Matches 195; Conservative 2; Mismatches 9; Indels 0; Gaps 0;

QY 112 MGGKWSKSSVVGWPTVTRMRRAEPAADGGAASRDLEKKGAISSNTAATNAACAWLEA 171
DB 1 MGGKWSKRIKGGWPAVRMRRAEPAADGGAASRDLEKKGAISSNTAATNAACAWLEA 60

QY 172 QEEVEGFPVTPQVPLRPMTYKAAVDLSHFLKEKGGLEGLIHSORRODILDWLWYHTQGY 231
DB 61 QEEVEGFPVTPQVPLRPMTYKAAVDLSHFLKEKGGLEGLIHSORRODILDWLWYHTQGY 120

QY 232 FPDQNTYTPGQVRYPLTFGWCYKLVPEPKVEEANKGENTSLHHPVSLHGMDPPER 291
DB 121 FPDQNTYTPGQVRYPLTFGWCYKLVPEPKVEEANKGENTSLHHPVSLHGMDPPER 180

QY 292 LEWRFDSRLAFHHVARELHPEYFKN 317
DB 181 LEWRFDSRLAFHHVARELHPEYFKN 206

RESULT 22
QY0595 PRELIMINARY; PRT; 206 AA.
AC Q90595;
DT 01-MAY-2000 (Tremblrel. 13, Created)
DT 01-MAY-2000 (Tremblrel. 13, Last sequence update)
DT 01-DEC-2001 (Tremblrel. 19, Last annotation update)
DE NEGATIVE FACTOR (F-PROTEIN) (27 KDA PROTEIN).
GN NEF.
OS Human immunodeficiency virus type 1.
OC Viruses; Retroid viruses; Retroviridae; Lentivirus.
OX NCBI_TaxID=11676;
RN [1]
RC STRAIN=8-IT;
RA Visco Comandini U., Yun Z., Paganelli R., Orlandi P., Salotti A.,
RA Johansson B., Vahine A., Sonnerborg A.;
RT "Hiv-1 nef mutations and clinical long-term non progression: a
RT molecular epidemiology study.";
RL Submitted (JUN-1997) to the EMBL/GenBank/DBJ databases.
CC -1- SIMILARITY: NEF HAS GTPASE, GTP-BINDING AND AUTOPHOSPHORYLATING
CC ACTIVITIES, IT SEEMS TO DOWN-REGULATE THE CD4(T4) ANTIGEN.
DR EMBL: AF011470; AAD01448.1; -
DR HSP: P03406; IEFN.
DR InterPro: IPR001558; F-protein.
DR Pfam: PF00469; F-protein; 1.
KW AIDS; Gtp-binding; Myristate.
SQ SEQUENCE 206 AA; 23681 MW; 1C891C3AD230C6A8 CRC64;

Query Match 47.3%; Score 1071; DB 15; Length 208;
Best Local Similarity 94.2%; Pred. No. 1.7e-78;
Matches 196; Conservative 3; Mismatches 7; Indels 2; Gaps 1;

QY 112 MGGKWSKSSVVGWPTVTRMRRAEPA--ADGGAASRDLEKKGAISSNTAATNAACAWL 169
DB 1 MGGKWSKRIKGGWPAVRMRRAEPAAXGAVSRDLEKKGAISSNTAATNAACAWL 60

QY 170 EAQEEVEGFPVTPQVPLRPMTYKAAVDLSHFLKEKGGLEGLIHSORRODILDWLWYHTQ 229
DB 61 EAQEEVEGFPVTPQVPLRPMTYKAAVDLSHFLKEKGGLEGLIHSORRODILDWLWYHTQ 120

QY 230 GYFPDQNTYTPGQVRYPLTFGWCYKLVPEPKVEEANKGENTSLHHPVSLHGMDPPER 289
DB 121 GYFPDQNTYTPGQVRYPLTFGWCYKLVPEPKVEEANKGENTSLHHPVSLHGMDPPER 180

QY 290 EVLEWRFDSRLAFHHVARELHPEYFKN 317
DB 181 KYLEWRFDSRLAFHHVARELHPEYFKN 208

RESULT 24
QY07X3 PRELIMINARY; PRT; 206 AA.
AC Q9W7X3;
DT 01-NOV-1999 (Tremblrel. 12, Created)
DT 01-NOV-1999 (Tremblrel. 12, Last sequence update)
DT 01-DEC-2001 (Tremblrel. 19, Last annotation update)
DE NEGATIVE FACTOR (F-PROTEIN) (27 KDA PROTEIN).
GN NEF.
OS Human immunodeficiency virus type 1.
OC Viruses; Retroid viruses; Retroviridae; Lentivirus.
OX NCBI_TaxID=11676;
RN [1]
```

```

QY 232 FPDQNTYTPGQVRYPLTFGWCYKLVPEPKVEEANKGENTSLHHPVSLHGMDPPER 291
DB 121 FPDQNTYTPGQVRYPLTFGWCYKLVPEPKVEEANKGENTSLHHPVSLHGMDPPER 180

QY 292 LEWRFDSRLAFHHVARELHPEYFKN 317
DB 181 LEWRFDSRLAFHHVARELHPEYFKN 206

RESULT 23
QY0YU9 PRELIMINARY; PRT; 208 AA.
AC Q9YU9;
DT 01-MAY-1999 (Tremblrel. 10, Created)
DT 01-MAY-1999 (Tremblrel. 10, Last sequence update)
DT 01-DEC-2001 (Tremblrel. 19, Last annotation update)
DE NEGATIVE FACTOR (F-PROTEIN) (27 KDA PROTEIN).
GN NEF.
OS Human immunodeficiency virus type 1.
OC Viruses; Retroid viruses; Retroviridae; Lentivirus.
OX NCBI_TaxID=11676;
RN [1]
RC STRAIN=8-IT;
RA Visco Comandini U., Yun Z., Paganelli R., Orlandi P., Salotti A.,
RA Johansson B., Vahine A., Sonnerborg A.;
RT "Hiv-1 nef mutations and clinical long-term non progression: a
RT molecular epidemiology study.";
RL Submitted (JUN-1997) to the EMBL/GenBank/DBJ databases.
CC -1- SIMILARITY: NEF HAS GTPASE, GTP-BINDING AND AUTOPHOSPHORYLATING
CC ACTIVITIES, IT SEEMS TO DOWN-REGULATE THE CD4(T4) ANTIGEN.
DR EMBL: AF011470; AAD01448.1; -
DR HSP: P03406; IEFN.
DR InterPro: IPR001558; F-protein.
DR Pfam: PF00469; F-protein; 1.
KW AIDS; Gtp-binding; Myristate.
SQ SEQUENCE 208 AA; 23681 MW; 1C891C3AD230C6A8 CRC64;

Query Match 47.3%; Score 1071; DB 15; Length 208;
Best Local Similarity 94.2%; Pred. No. 1.7e-78;
Matches 196; Conservative 3; Mismatches 7; Indels 2; Gaps 1;

QY 112 MGGKWSKSSVVGWPTVTRMRRAEPA--ADGGAASRDLEKKGAISSNTAATNAACAWL 169
DB 1 MGGKWSKRIKGGWPAVRMRRAEPAAXGAVSRDLEKKGAISSNTAATNAACAWL 60

QY 170 EAQEEVEGFPVTPQVPLRPMTYKAAVDLSHFLKEKGGLEGLIHSORRODILDWLWYHTQ 229
DB 61 EAQEEVEGFPVTPQVPLRPMTYKAAVDLSHFLKEKGGLEGLIHSORRODILDWLWYHTQ 120

QY 230 GYFPDQNTYTPGQVRYPLTFGWCYKLVPEPKVEEANKGENTSLHHPVSLHGMDPPER 289
DB 121 GYFPDQNTYTPGQVRYPLTFGWCYKLVPEPKVEEANKGENTSLHHPVSLHGMDPPER 180

QY 290 EVLEWRFDSRLAFHHVARELHPEYFKN 317
DB 181 KYLEWRFDSRLAFHHVARELHPEYFKN 208

RESULT 24
QY07X3 PRELIMINARY; PRT; 206 AA.
AC Q9W7X3;
DT 01-NOV-1999 (Tremblrel. 12, Created)
DT 01-NOV-1999 (Tremblrel. 12, Last sequence update)
DT 01-DEC-2001 (Tremblrel. 19, Last annotation update)
DE NEGATIVE FACTOR (F-PROTEIN) (27 KDA PROTEIN).
GN NEF.
OS Human immunodeficiency virus type 1.
OC Viruses; Retroid viruses; Retroviridae; Lentivirus.
OX NCBI_TaxID=11676;
RN [1]
```

RP SEQUENCE FROM N.A.  
 RC STRAIN-PATIENT 27;  
 RA McPhee D.A., Greenway A.L., Holloway G., Smith K., Deacon N.,  
 RA Pemberton L., Brew B.J.;  
 RT "Anomalies in Nef expression within the central nervous system of HIV-  
 RT 1 positive individuals/AIDS patients with or without AIDS dementia  
 RT complex";  
 RL J. Neurovirol. 4:0-0(1998).  
 CC -1- SIMILARITY: NEF HAS GTPASE, GTP-BINDING AND AUTOPHOSPHORYLATING  
 CC ACTIVITIES, IT SEEMS TO DOWN-REGULATE THE CD4(T4) ANTIGEN.  
 DR EMBL: AF064675; AAC18377.1; -;  
 DR HSSP: P03406; IEFN.  
 DR InterPro: IPR001558; F-protein.  
 DR Pfam: PF00469; F-protein; 1.  
 KW AIDS; GTP-binding; Myristate.  
 SQ SEQUENCE 206 AA; 23499 MW; DEE21CAF05891D6B CRC64;

Query Match 47.2%; Score 1068; DB 15; Length 206;  
 Best Local Similarity 91.7%; Pred. No. 3e-78; 3; Indels 0; Gaps 0;  
 Matches 189; Conservative 14; Mismatches 14; Indels 0; Gaps 0;  
 QY 112 MGGKWSKSSVVGWPTVRERMRRAEPAADGVGAASRDLEKKGAISSNTAATNAACAWLEA 171  
 Db 1 MGGKWSKSSVVGWPTVRERMRRAEPAADGVGAASRDLEKKGAISSNTAATNAACAWLEA 60  
 QY 172 QEEVEVGFPVTPQVPLRPMTYKAAVDLSHFLKEKGGLEGLIHSORRQDILDWIYHTOGY 231  
 Db 61 QEEVEVGFPVTPQVPLRPMTYKAAVDLSHFLKEKGGLEGLIHSORRQDILDWIYHTOGY 120  
 QY 232 FPDWQNTYTPGCVRYPLTFGWCYKLVPEPKVEEANKGENTSLHHPVSLHGMDDPEREV 291  
 Db 121 FPDWQNTYTPGCVRYPLTFGWCYKLVPEPKVEEANKGENTSLHHPVSLHGMDDPEREV 180  
 QY 292 LEWRDLSRALPHHHVARELHPEYFKNC 317  
 Db 181 LEWRDLSRALPHHHVARELHPEYFKNC 206

RESULT 25  
 Q9DQ01 PRELIMINARY; PRT: 206 AA.  
 AC Q9DQ01;  
 DT 01-MAR-2001 (TrEMBLrel. 16, Created)  
 DT 01-MAR-2001 (TrEMBLrel. 16, Last sequence update)  
 DT 01-DEC-2001 (TrEMBLrel. 19, Last annotation update)  
 DE NEGATIVE FACTOR (F-PROTEIN) (27 KDA PROTEIN).  
 GN NEF.  
 OS Human immunodeficiency virus type 1.  
 OC Viruses; Retrovirdae; Lentiviridae; Lentivirus.  
 OX NCBI\_TaxID=11676;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN=03U03;  
 RA MEDLINE=21002575; PubMed=11118071;  
 RA Geffin R., Wolf D., Muller R., Hill M.D., Stellwag E., Freitag M.,  
 RA Saks G., Scott G.B., Baur A.S.;  
 RT "Functional and structural defects in HIV-1 nef genes derived from  
 RT pediatric long-term survivors";  
 RL AIDS Res. Hum. Retroviruses 16:1855-1868(2000).  
 CC -1- SIMILARITY: NEF HAS GTPASE, GTP-BINDING AND AUTOPHOSPHORYLATING  
 CC ACTIVITIES, IT SEEMS TO DOWN-REGULATE THE CD4(T4) ANTIGEN.  
 DR EMBL: AF252911; AAG34580.1; -;  
 DR HSSP: P03406; IEFN.  
 DR InterPro: IPR001558; F-protein.  
 DR Pfam: PF00469; F-protein; 1.  
 KW AIDS; GTP-binding; Myristate.  
 SQ SEQUENCE 206 AA; 23464 MW; B72E0AA736374B64 CRC64;

Query Match 47.0%; Score 1065; DB 15; Length 206;  
 Best Local Similarity 93.2%; Pred. No. 5.2e-78;  
 Matches 192; Conservative 9; Mismatches 5; Indels 0; Gaps 0;

QY 112 MGGKWSKSSVVGWPTVRERMRRAEPAADGVGAASRDLEKKGAISSNTAATNAACAWLEA 171  
 Db 1 MGGKWSKSSVVGWPTVRERMRRAEPAADGVGAASRDLEKKGAISSNTAATNAACAWLEA 60  
 QY 172 QEEVEVGFPVTPQVPLRPMTYKAAVDLSHFLKEKGGLEGLIHSORRQDILDWIYHTOGY 231  
 Db 61 QEEVEVGFPVTPQVPLRPMTYKAAVDLSHFLKEKGGLEGLIHSORRQDILDWIYHTOGY 120  
 QY 232 FPDWQNTYTPGCVRYPLTFGWCYKLVPEPKVEEANKGENTSLHHPVSLHGMDDPEREV 291  
 Db 121 FPDWQNTYTPGCVRYPLTFGWCYKLVPEPKVEEANKGENTSLHHPVSLHGMDDPEREV 180  
 QY 292 LEWRDLSRALPHHHVARELHPEYFKNC 317  
 Db 181 LEWRDLSRALPHHHVARELHPEYFKNC 206

RESULT 26  
 Q9W7U0 PRELIMINARY; PRT: 206 AA.  
 AC Q9W7U0;  
 DT 01-NOV-1999 (TrEMBLrel. 12, Created)  
 DT 01-NOV-1999 (TrEMBLrel. 12, Last sequence update)  
 DT 01-DEC-2001 (TrEMBLrel. 19, Last annotation update)  
 DE NEGATIVE FACTOR (F-PROTEIN) (27 KDA PROTEIN).  
 GN NEF.  
 OS Human immunodeficiency virus type 1.  
 OC Viruses; Retrovirdae; Lentiviridae; Lentivirus.  
 OX NCBI\_TaxID=11676;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN=PATIENT 27;  
 RA McPhee D.A., Greenway A.L., Holloway G., Smith K., Deacon N.,  
 RA Pemberton L., Brew B.J.;  
 RT "Anomalies in Nef expression within the central nervous system of HIV-  
 RT 1 positive individuals/AIDS patients with or without AIDS dementia  
 RT complex";  
 RL J. Neurovirol. 4:0-0(1998).  
 CC -1- SIMILARITY: NEF HAS GTPASE, GTP-BINDING AND AUTOPHOSPHORYLATING  
 CC ACTIVITIES, IT SEEMS TO DOWN-REGULATE THE CD4(T4) ANTIGEN.  
 DR EMBL: AF064675; AAC18376.1; -;  
 DR EMBL: AF064674; AAC18375.1; -;  
 DR HSSP: P03406; IEFN.  
 DR InterPro: IPR001558; F-protein.  
 DR Pfam: PF00469; F-protein; 1.  
 KW AIDS; GTP-binding; Myristate.  
 SQ SEQUENCE 206 AA; 23514 MW; DC681CAF05891D6B CRC64;

Query Match 47.0%; Score 1064; DB 15; Length 206;  
 Best Local Similarity 91.7%; Pred. No. 6.2e-78;  
 Matches 189; Conservative 13; Mismatches 4; Indels 0; Gaps 0;

QY 112 MGGKWSKSSVVGWPTVRERMRRAEPAADGVGAASRDLEKKGAISSNTAATNAACAWLEA 171  
 Db 1 MGGKWSKSSVVGWPTVRERMRRAEPAADGVGAASRDLEKKGAISSNTAATNAACAWLEA 60  
 QY 172 QEEVEVGFPVTPQVPLRPMTYKAAVDLSHFLKEKGGLEGLIHSORRQDILDWIYHTOGY 231  
 Db 61 QEEVEVGFPVTPQVPLRPMTYKAAVDLSHFLKEKGGLEGLIHSORRQDILDWIYHTOGY 120  
 QY 232 FPDWQNTYTPGCVRYPLTFGWCYKLVPEPKVEEANKGENTSLHHPVSLHGMDDPEREV 291  
 Db 121 FPDWQNTYTPGCVRYPLTFGWCYKLVPEPKVEEANKGENTSLHHPVSLHGMDDPEREV 180  
 QY 292 LEWRDLSRALPHHHVARELHPEYFKNC 317  
 Db 181 LEWRDLSRALPHHHVARELHPEYFKNC 206

RESULT 27  
 Q93010

```
ID O93010 PRELIMINARY; PRT; 206 AA.
AC O93010;
DT 01-NOV-1998 (TReMBLrel. 08, Created)
DT 01-NOV-1998 (TReMBLrel. 08, Last sequence update)
DT 01-DEC-2001 (TReMBLrel. 19, Last annotation update)
DE NEGATIVE FACTOR (F-PROTEIN) (27 KDA PROTEIN).
GN NEF.
OS Human immunodeficiency virus type 1.
OC Viruses; Retrovirdae; Retroviridae; Lentivirus.
OX NCBI_TaxID=11676;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN-PATIENT 9;
RX MEDLINE=98097260; PubMed=9436760;
RA Kang M.R., Cho Y.K., Chun J., Kim Y.B., Lee I., Lee H.J., Kim S.H.,
RA Kim Y.K., Yoon K., Yang J.M., Kim J.M., Shin Y.O., Kang C., Lee J.S.,
RA Choi K.W., Kim D.G., Fitch W.M., Kim S.;
RT "Phylogenetic analysis of the nef gene reveals a distinctive
RT monophyletic clade in Korean HIV-1 cases.";
RL J. Acquir. Immune Defic. Syndr. Hum. Retrovirol. 17:58-68(1998).
CC -1- SIMILARITY: NEF HAS GTPASE, GTP-BINDING AND AUTOPHOSPHORYLATING
CC ACTIVITIES, IT SEEMS TO DOWN-REGULATE THE CD4(T4) ANTIGEN.
DR EMBL; AF063922; AAC17893.1; -
DR HSP; P03406; IEFN.
DR InterPro; IPR001558; F-protein.
DR Pfam; PF00469; F-protein; 1.
DR AIDS; GTP-binding; Myristate.
KW SEQUENCE 206 AA; 23348 MW; 4D5AEDF55FAE93E3 CRC64;
SQ
```

```
Query Match 46.9%; Score 1062; DB 15; Length 206;
Best Local Similarity 94.68; Pred. No. 9e-78; Indels 0; Gaps 0;
Matches 194; Conservative 3; Mismatches 8; Indels 0; Gaps 0;

Qy 112 MGGKWSKSSVVGWPTVRMRRAEPAADVGGAASRDLEKKGAISSNTAATNAACAWLEA 171
Db 1 MGGKWSKSSVVGWPTVRMRRAEPAADVGGAASRDLEKKGAISSNTAATNAACAWLEA 60
Qy 172 QEEVEVGPVTPQVPLRPMTYKAAVDLSHFLKKGGLGSLHSORRQDILDWIYHTQGY 231
Db 61 QEEVEVGPVTPQVPLRPMTYKAAVDLSHFLKKGGLGSLHSORRQDILDWIYHTQGY 120
Qy 232 FPDQNTYTPGCVRYPLTFGCYKLVPEPKVEANKGENTSLHPVSLHGMDPPEV 291
Db 121 FPDQNTYTPGCVRYPLTFGCYKLVPEPKVEANKGENTSLHPVSLHGMDPPEV 180
Qy 292 LEWRFSRLAFHHVARELHPEYFKN 316
Db 181 LEWRFSRLAFHHVARELHPEYKDC 206
```

```
RESULT 28
O93012
ID O93012 PRELIMINARY; PRT; 206 AA.
AC O93012;
DT 01-NOV-1998 (TReMBLrel. 08, Created)
DT 01-NOV-1998 (TReMBLrel. 08, Last sequence update)
DT 01-DEC-2001 (TReMBLrel. 19, Last annotation update)
DE NEGATIVE FACTOR (F-PROTEIN) (27 KDA PROTEIN).
GN NEF.
OS Human immunodeficiency virus type 1.
OC Viruses; Retrovirdae; Retroviridae; Lentivirus.
OX NCBI_TaxID=11676;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN-PATIENT 27;
RA McPhee D.A., Greenway A.L., Holloway G., Smith K., Deacon N.,
RA Pemberton L., Brew B.J.;
RT "Anomalies in Nef expression within the central nervous system of HIV-
RT 1 positive individuals/AIDS patients with or without AIDS dementia
RT complex.";
RL J. Neurovirol. 4:0-0(1998).
CC -1- SIMILARITY: NEF HAS GTPASE, GTP-BINDING AND AUTOPHOSPHORYLATING
```

```
CC ACTIVITIES, IT SEEMS TO DOWN-REGULATE THE CD4(T4) ANTIGEN.
DR EMBL; AF064673; AAC18374.1; -
DR HSP; P03406; IEFN.
DR InterPro; IPR001558; F-protein.
DR Pfam; PF00469; F-protein; 1.
KW AIDS; GTP-binding; Myristate.
SQ SEQUENCE 206 AA; 23472 MW; 2255A447EC8E5456 CRC64;

Query Match 46.9%; Score 1061; DB 15; Length 206;
Best Local Similarity 91.7%; Pred. No. 1.1e-77;
Matches 189; Conservative 12; Mismatches 5; Indels 0; Gaps 0;

Qy 112 MGGKWSKSSVVGWPTVRMRRAEPAADVGGAASRDLEKKGAISSNTAATNAACAWLEA 171
Db 1 MGGKWSKSSVVGWPTVRMRRAEPAADVGGAASRDLEKKGAISSNTAATNAACAWLEA 60
Qy 172 QEEVEVGPVTPQVPLRPMTYKAAVDLSHFLKKGGLGSLHSORRQDILDWIYHTQGY 231
Db 61 QEEVEVGPVTPQVPLRPMTYKAAVDLSHFLKKGGLGSLHSORRQDILDWIYHTQGY 120
Qy 232 FPDQNTYTPGCVRYPLTFGCYKLVPEPKVEANKGENTSLHPVSLHGMDPPEV 291
Db 121 FPDQNTYTPGCVRYPLTFGCYKLVPEPKVEANKGENTSLHPVSLHGMDPPEV 180
Qy 292 LEWRFSRLAFHHVARELHPEYFKN 317
Db 181 LEWRFSRLAFHHVAREKHPYKDC 206
```

```
RESULT 29
O9Q597
ID O9Q597 PRELIMINARY; PRT; 206 AA.
AC O9Q597;
DT 01-MAY-2000 (TReMBLrel. 13, Created)
DT 01-MAY-2000 (TReMBLrel. 13, Last sequence update)
DT 01-DEC-2001 (TReMBLrel. 19, Last annotation update)
DE NEGATIVE FACTOR (F-PROTEIN) (27 KDA PROTEIN).
GN NEF.
OS Human immunodeficiency virus type 1.
OC Viruses; Retrovirdae; Retroviridae; Lentivirus.
OX NCBI_TaxID=11676;
RN [1]
RP SEQUENCE FROM N.A.
RA Wang B., Saksena N.K.;
RT "HIV-1 Strains from a cohort of American subjects reveal the presence
RT of a V2 region extension unique to slow progressors and non-
RT progressors.";
RL AIDS 0:0-0(2000).
CC -1- SIMILARITY: NEF HAS GTPASE, GTP-BINDING AND AUTOPHOSPHORYLATING
CC ACTIVITIES, IT SEEMS TO DOWN-REGULATE THE CD4(T4) ANTIGEN.
DR EMBL; AF203196; AAF25318.1; -
DR HSP; P03406; IEFN.
DR InterPro; IPR001558; F-protein.
DR Pfam; PF00469; F-protein; 1.
KW AIDS; GTP-binding; Myristate.
SQ SEQUENCE 206 AA; 23599 MW; 9C887A58541A9905 CRC64;

Query Match 46.8%; Score 1060; DB 15; Length 206;
Best Local Similarity 92.7%; Pred. No. 1.3e-77;
Matches 191; Conservative 7; Mismatches 8; Indels 0; Gaps 0;

Qy 112 MGGKWSKSSVVGWPTVRMRRAEPAADVGGAASRDLEKKGAISSNTAATNAACAWLEA 171
Db 1 MGGKWSKSSVVGWPTVRMRRAEPAADVGGAASRDLEKKGAISSNTAATNAACAWLEA 60
Qy 172 QEEVEVGPVTPQVPLRPMTYKAAVDLSHFLKKGGLGSLHSORRQDILDWIYHTQGY 231
Db 61 QEEVEVGPVTPQVPLRPMTYKAAVDLSHFLKKGGLGSLHSORRQDILDWIYHTQGY 120
Qy 232 FPDQNTYTPGCVRYPLTFGCYKLVPEPKVEANKGENTSLHPVSLHGMDPPEV 291
Db 181 LEWRFSRLAFHHVAREKHPYKDC 206
```

Db 121 FPDQNYTGPGRVPLTGWCKLVPEQKVKANEKNTSLHPSLHGMDDEPERV 180  
QY 292 LEWRDLSRLAFHHVARELHPEYFKNC 317  
Db 181 LEWRDLSRLAFHHVARELHPEYFKNC 206

## RESULT 30

Q9W7X2  
ID Q9W7X2 PRELIMINARY: PRT: 206 AA.  
AC Q9W7X2;  
DT 01-NOV-1999 (TREMELrel. 12, Created)  
DT 01-NOV-1999 (TREMELrel. 12, Last sequence update)  
DT 01-DEC-2001 (TREMELrel. 19, Last annotation update)  
DE NEGATIVE FACTOR (F-PROTEIN) (27 KDA PROTEIN).  
GN NEF.  
OS Human immunodeficiency virus type 1.  
OC Viruses; Retrovirdae; Retroviridae; Lentivirus.  
OX NCBI\_TaxID=11676;  
RN (1)  
RP SEQUENCE FROM N.A.  
RC STRAIN-PATIENT 27;  
RA McPhee D.A., Greenway A.L., Holloway G., Smith K., Deacon N.,  
Pemberton L., Brew B.J.;  
RT "Anomalies in Nef expression within the central nervous system of HIV-  
RT 1 positive individuals/AIDS patients with or without AIDS dementia  
RT complex."  
RL J. Neurovirol. 4:0-0(1998).  
CC -I- SIMILARITY: NEF HAS GTPASE, GTP-BINDING AND AUTOPHOSPHORYLATING  
CC ACTIVITIES, IT SEEMS TO DOWN-REGULATE THE CD4(T4) ANTIGEN.  
DR EMBL; AF064677; AAC18378.1; -  
DR HSP; P03406; 1EFN.  
DR InterPro: IPR001558; F-protein.  
DR Pfam: PF00469; F-protein; 1.  
DR AIDS; GTP-binding; Myristate.  
KW AIDS; GTP-binding; Myristate.  
SQ SEQUENCE 206 AA; 23442 MW; DC6A96AF05891D6B CRC64;

Query Match 46.7%; Score 1057; DB 15; Length 206;  
Best Local Similarity 91.3%; Pred. No. 2.3e-77;  
Matches 188; Conservative 13; Mismatches 5; Indels 0; Gaps 0;  
QY 112 MGCKKSKSSVVGWPTVRRMRRAEPAADGVGAASRDLEKHGAISSNTAATNAACAWLEA 171  
Db 1 MGCKKSKSSVVGWPTVRRMRRAEPAADGVGAASRDLEKHGAISSNTAATNAACAWLEA 60  
QY 172 QEEVEGVFPVTPQVPLRPMTYKAAVDLSHFLKEKGGLEGLIHSQRQDILDLYHTQGY 231  
Db 61 QKEEEGVFPVTPQVPLRPMTYKAAVDLSHFLKEKGGLEGLIHSQRQDILDLYHTQGY 120  
QY 232 FPDQNYTGPGRVPLTGWCKLVPEQKVKANEKNTSLHPSLHGMDDEPERV 291  
Db 121 FPDQNYTGPGRVPLTGWCKLVPEQKVKANEKNTSLHPSLHGMDDEPERV 180  
QY 292 LEWRDLSRLAFHHVARELHPEYFKNC 317  
Db 181 LEWRDLSRLAFHHVARELHPEYFKNC 206

Search completed: August 26, 2002, 08:14:57  
Job time: 385 sec

GenCore version 4.5  
Copyright (c) 1993 - 2000 Compugen Ltd.

OM protein - protein search, using sw model

Run on: August 26, 2002, 08:12:30 ; Search time 99.16 Seconds  
(without alignments)  
106.414 Million cell updates/sec

Title: US-09-509-239-23  
Perfect score: 552  
Sequence: 1 MPVDPRLFPKHPGSPKT.....QSKGPTGPKETSGHHHHH 95

Scoring table: BLOSUM62  
Gapop 10.0 , Gapext 0.5

Searched: 747574 seqs, 111073796 residues

Total number of hits satisfying chosen parameters: 747574

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :  
1: /SIDS1/gcgdata/hold-geneseq/geneseq-emb1/AA1980.DAT:  
2: /SIDS1/gcgdata/hold-geneseq/geneseq-emb1/AA1981.DAT:  
3: /SIDS1/gcgdata/hold-geneseq/geneseq-emb1/AA1982.DAT:  
4: /SIDS1/gcgdata/hold-geneseq/geneseq-emb1/AA1983.DAT:  
5: /SIDS1/gcgdata/hold-geneseq/geneseq-emb1/AA1984.DAT:  
6: /SIDS1/gcgdata/hold-geneseq/geneseq-emb1/AA1985.DAT:  
7: /SIDS1/gcgdata/hold-geneseq/geneseq-emb1/AA1986.DAT:  
8: /SIDS1/gcgdata/hold-geneseq/geneseq-emb1/AA1987.DAT:  
9: /SIDS1/gcgdata/hold-geneseq/geneseq-emb1/AA1988.DAT:  
10: /SIDS1/gcgdata/hold-geneseq/geneseq-emb1/AA1989.DAT:  
11: /SIDS1/gcgdata/hold-geneseq/geneseq-emb1/AA1990.DAT:  
12: /SIDS1/gcgdata/hold-geneseq/geneseq-emb1/AA1991.DAT:  
13: /SIDS1/gcgdata/hold-geneseq/geneseq-emb1/AA1992.DAT:  
14: /SIDS1/gcgdata/hold-geneseq/geneseq-emb1/AA1993.DAT:  
15: /SIDS1/gcgdata/hold-geneseq/geneseq-emb1/AA1994.DAT:  
16: /SIDS1/gcgdata/hold-geneseq/geneseq-emb1/AA1995.DAT:  
17: /SIDS1/gcgdata/hold-geneseq/geneseq-emb1/AA1996.DAT:  
18: /SIDS1/gcgdata/hold-geneseq/geneseq-emb1/AA1997.DAT:  
19: /SIDS1/gcgdata/hold-geneseq/geneseq-emb1/AA1998.DAT:  
20: /SIDS1/gcgdata/hold-geneseq/geneseq-emb1/AA1999.DAT:  
21: /SIDS1/gcgdata/hold-geneseq/geneseq-emb1/AA2000.DAT:  
22: /SIDS1/gcgdata/hold-geneseq/geneseq-emb1/AA2001.DAT:

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

# SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	552	100.0	95	20	AA1980.DAT: A representative H
2	552	100.0	95	22	AA1981.DAT: Amino acid sequenc
3	547	99.1	302	20	AA1982.DAT: A representative H
4	547	99.1	302	22	AA1983.DAT: Amino acid sequenc
5	541	98.0	95	20	AA1984.DAT: A representative H
6	541	98.0	95	22	AA1985.DAT: Amino acid sequenc
7	536	97.1	302	20	AA1986.DAT: A representative H
8	536	97.1	302	22	AA1987.DAT: Amino acid sequenc
9	536	97.1	411	20	AA1988.DAT: A representative L
10	536	97.1	411	22	AA1989.DAT: Amino acid sequenc
11	536	97.1	413	20	AA1990.DAT: A representative L

12	536	97.1	413	22	AAG63237	Amino acid sequenc
13	479	86.8	86	20	AA122960	Lys41 mutant tat a
14	478	86.6	86	8	AAP70598	HIV virus recombin
15	478	86.6	86	10	AAP91903	Tat protein derive
16	478	86.6	86	10	AAP91905	CDNA for tat prote
17	478	86.6	86	10	AAP93140	Synthetic Tat prot
18	478	86.6	86	14	AAR8618	Sequence of the HI
19	478	86.6	86	17	AAR86625	HIV Tat. Human im
20	478	86.6	86	18	AAW31202	Human Immunodefici
21	478	86.6	86	18	AAW26443	HIV tat protein.
22	478	86.6	86	19	AAR71385	Amino acid sequenc
23	478	86.6	86	19	AAR76148	HIV Type 1 Tat pro
24	478	86.6	86	21	AA996881	Human immunodefici
25	478	86.6	86	22	AAB85996	Amino acid sequenc
26	478	86.6	87	12	AAR13379	TAT protein. Synt
27	466	84.4	86	20	AA242959	Cys22 mutant tat a
28	462	83.7	86	11	AAR02327	Sequence of the tr
29	462	83.7	86	16	AAR78742	Wild type Tat HIV
30	457.5	82.9	83	20	AA222961	Lys41-RGD-delta mu
31	456.5	82.7	83	20	AA222962	RGD-delta mutant t
32	452	81.9	83	20	AA222958	Wild type Tat amin
33	451	81.7	86	15	AAR48966	Diversified HIV-1
34	451	81.7	86	15	AAR57346	Peptide fragment o
35	448	81.2	86	21	AAB10050	HIV-1 tat protein.
36	446	80.8	86	15	AAR48964	Full length Tat.
37	446	80.8	86	22	AG644257	Human protein: SEQ
38	442	80.1	86	22	AAB46679	HIV-1 Tat peptide
39	442	80.1	86	22	AAB46685	HIV-1 Tat peptide
40	433	78.4	86	16	AAR77542	Trans-dominant var
41	430	77.9	101	21	AAB14224	HIV Tat-SF162. Hu
42	418	75.7	102	21	AAB14225	HIV Ta_Cys22_SF16
43	411	74.5	101	12	AAR12259	HIV-1 strain Oyi T
44	411	74.5	101	21	AAB30514	Sequence of the ta
45	408	73.9	72	7	AAP60699	Trans-acting trans

## ALIGNMENTS

RESULT 1  
AA1980.DAT: Amino acid sequenc  
ID AA1980.DAT: Amino acid sequenc; Protein; 95 AA.  
XX  
XX AA1981.DAT: HIV virus recombin  
AC AA1982.DAT: Tat protein derive  
DT 09-JUL-1999 (first entry)  
DE A representative HIV-1 mutant Nef-His protein.  
XX HIV nef gene; fusion protein; HIV nef gene; Nef protein; Tat protein;  
KW vaccine; HIV infection; protein D.  
XX Synthetic.  
OS Human immunodeficiency virus type 1.  
XX  
PN WO9916884-A1.  
XX  
PD 08-APR-1999.  
PF 17-SEP-1998; 98WO-EP06040.  
XX  
PR 26-SEP-1997; 97GB-0020585.  
XX (SMK ) SMITHKLINE BEECHAM BIOLOGICALS.  
PA Bruck C, Godart SAG, Marchand M;  
PI WPI; 1999-302282/25.  
DR N-PSDB; AA19835692.  
XX HIV Tat or Nef protein linked to a fusion partner  
PT Disclosure; Fig 2; 66pp; English.  
PS

XX CC The present sequence represents a representative HIV-1 mutant Nef-His  
 CC protein. The protein is used in the creation of the fusion proteins of  
 CC the invention, in conjunction with a fusion partner (e.g. protein D).  
 CC The specification also describes fusion proteins comprising HIV-1 Tat  
 CC protein. The fusion protein can be used in a vaccine to prevent HIV  
 CC infection.

XX SQ Sequence 95 AA;

Query Match 100.0%; Score 552; DB 20; Length 95;  
 Best Local Similarity 100.0%; Pred. No. 3.4e-50;  
 Matches 95; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MEPVDPRLPEPKHPSQKPTACTNCYCKKCCFHCQVCFITAAALGISYGRKKRRRRPPQ 60  
 Db 1 mepvdprlepwhpgsqptactncycckkccfchqvcfitaalgisygrkkrrrrppq 60

QY 61 GSOTHOVSLSKQPTSQSKGPTGPKETSGHHHHH 95  
 Db 61 gsqthqvsksqptsqskgptgpketsgghhhhh 95

RESULT 2

AAAG63238  
 ID AAG63238 standard; Protein; 95 AA.  
 XX  
 AC AAG63238;  
 DT 01-OCT-2001 (first entry)  
 XX  
 DE Amino acid sequence of a His-tagged mutant Tat protein of HIV.  
 XX  
 KW HIV; nef gene; tat gene; vaccine; gp120 gene; HIV viral load.  
 XX  
 OS Synthetic.  
 OS Human immunodeficiency virus.  
 XX  
 PN WO200154719-A2.  
 XX  
 PD 02-AUG-2001.  
 XX  
 PE 29-JAN-2001; 2001WO-EP00944.  
 XX  
 PR 31-JAN-2000; 2000GB-0002200.  
 PR 14-APR-2000; 2000GB-0009336.  
 PR 06-JUN-2000; 2000GB-0013806.  
 PR 28-JUN-2000; 2000WO-EP05998.  
 XX  
 PA (SMIK ) SMITHKLINE BEECHAM BIOLOGICALS.  
 XX  
 PI Voss G;  
 XX  
 WPI; 2001-476172/51.  
 DR N-PSDB; AAH42882.  
 XX  
 PT New use of a human immunodeficiency virus (HIV) Tat, or Nef, or Nef  
 PT linked to Tat (Nef-Tat) protein or polynucleotide and an HIV gp120  
 PT protein or polynucleotide for the manufacture of a vaccine .  
 XX  
 PS Disclosure; Fig 1; 90pp; English.  
 XX  
 CC The present sequence represents a His-tagged mutant Tat protein of HIV.  
 CC The protein is expressed in the yeast Pichia pastoris, and is used to  
 CC produce the vaccine of the invention. The specification describes  
 CC the use of HIV Tat, HIV Nef, or Nef-Tat; and HIV gp120 in the  
 CC manufacture of a vaccine. The vaccine is used for the prophylactic or  
 CC therapeutic immunization of humans against HIV. Tat, Nef or Nef-Tat act  
 CC in synergy with gp120 in the treatment and prevention of HIV. The  
 CC vaccine reduces the HIV viral load in HIV infected humans and results  
 CC in a maintenance of CD4+ levels over those levels found in the absence  
 CC of vaccination with HIV Tat, Nef or Nef-Tat and HIV gp120.

XX SQ Sequence 95 AA;

Query Match 100.0%; Score 552; DB 22; Length 95;  
 Best Local Similarity 100.0%; Pred. No. 3.4e-50;  
 Matches 95; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MEPVDPRLPEPKHPSQKPTACTNCYCKKCCFHCQVCFITAAALGISYGRKKRRRRPPQ 60  
 Db 1 mepvdprlepwhpgsqptactncycckkccfchqvcfitaalgisygrkkrrrrppq 60

QY 61 GSOTHOVSLSKQPTSQSKGPTGPKETSGHHHHH 95  
 Db 61 gsqthqvsksqptsqskgptgpketsgghhhhh 95

RESULT 3

AAAY02357  
 ID AAY02357 standard; Protein; 302 AA.  
 XX  
 AC AAY02357;  
 DT 09-JUL-1999 (first entry)  
 XX  
 DE A representative HIV-1 mutant Tat-His protein.  
 XX  
 KW HIV nef gene; fusion protein; HIV nef gene; Nef protein; Tat protein;  
 KW vaccine; HIV infection; protein D.  
 XX  
 OS Synthetic.  
 OS Human immunodeficiency virus type 1.  
 XX  
 PN WO9916884-A1.  
 XX  
 PD 08-APR-1999.  
 XX  
 PE 17-SEP-1998; 98WO-EP06040.  
 XX  
 PR 26-SEP-1997; 97GB-0020585.  
 XX  
 PA (SMIK ) SMITHKLINE BEECHAM BIOLOGICALS.  
 XX  
 PI Bruck C, Godart SAG, Marchand M;  
 XX  
 WPI; 1999-302282/25.  
 DR N-PSDB; AAX35693.  
 XX  
 PT HIV Tat or Nef protein linked to a fusion partner  
 XX  
 PS Disclosure; Fig 2; 66pp; English.  
 XX  
 CC The present sequence represents a representative HIV-1 mutant Tat-His  
 CC protein. The protein is used in the creation of the fusion proteins of  
 CC the invention, in conjunction with a fusion partner (e.g. protein D).  
 CC The specification also describes fusion proteins comprising HIV-1 Nef  
 CC protein. The fusion protein can be used in a vaccine to prevent HIV  
 CC infection.

XX SQ Sequence 302 AA;

Query Match 99.1%; Score 547; DB 20; Length 302;  
 Best Local Similarity 100.0%; Pred. No. 3.8e-49;  
 Matches 94; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 2 EPVDPRLPEPKHPSQKPTACTNCYCKKCCFHCQVCFITAAALGISYGRKKRRRRPPQ 61  
 Db 209 epvdprlepwhpgsqptactncycckkccfchqvcfitaalgisygrkkrrrrppq 268

QY 62 SOTHOVSLSKQPTSQSKGPTGPKETSGHHHHH 95  
 Db 269 sqthqvsksqptsqskgptgpketsgghhhhh 302

RESULT 4  
AAG63239  
ID AAG63239 standard; Protein; 302 AA.

XX AC AAG63239;  
XX DT 01-OCT-2001 (first entry)  
XX DE Amino acid sequence of a His-tagged mutant His protein of HIV.  
XX KW HIV; nef gene; tat gene; vaccine; gp120 gene; HIV viral load.  
XX OS Synthetic.  
XX OS Human immunodeficiency virus.

XX PN WO200154719-A2.  
XX PD 02-AUG-2001.

XX PF 29-JAN-2001; 2001WO-EP00944.

XX PR 31-JAN-2000; 2000GB-0002200.

XX PR 14-APR-2000; 2000GB-0009336.

XX PR 06-JUN-2000; 2000GB-0013806.

XX PR 28-JUN-2000; 2000WO-EP05998.

XX PA (SMIK ) SMITHKLINE BEECHAM BIOLOGICALS.  
XX PI Voss G;  
XX XX

XX DR WPI; 2001-476172/51.

XX DR N-PSDB; AAH42883.

XX PT New use of a human immunodeficiency virus (HIV) Tat, or Nef, or Nef  
XX PT linked to Tat (Nef-Tat) protein or polynucleotide and an HIV gp120  
XX PT protein or polynucleotide for the manufacture of a vaccine -  
XX XX

XX PS Disclosure; Fig 1; 90pp; English.

XX CC The present sequence represents a His-tagged mutant His protein of HIV.  
XX CC The protein is expressed in the yeast Pichia pastoris, and is used to  
XX CC produce the vaccine of the invention. The specification describes  
XX CC the use of HIV Tat, HIV Nef, or Nef-Tat; and HIV gp120 in the  
XX CC manufacture of a vaccine. The vaccine is used for the prophylactic or  
XX CC therapeutic immunization of humans against HIV. Tat, Nef or Nef-Tat act  
XX CC in synergy with gp120 in the treatment and prevention of HIV. The  
XX CC vaccine reduces the HIV viral load in HIV infected humans and results  
XX CC in a maintenance of CD4+ levels over those levels found in the absence  
XX CC of vaccination with HIV Tat, Nef or Nef-Tat and HIV gp120.

XX SQ Sequence 302 AA;

Query Match 99.18; Score 547; DB 22; Length 302;  
Best Local Similarity 100.0%; Pred. No. 3.8e-49;  
Matches 94; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 2 EPVDPRLPEPWKHPGSOPTACTNCYCKKCCFHCQVCFITAAALGISYGRKKRRRRPPQG 61  
Db 209 epvdpriepwkhpgsqptactncycckkccfhcqvcfitaalgisyrkrrrrppqg 268

Qy 62 SQTHQVLSKQPTSQSGKEPTGPKETSGHHHHH 95  
Db 269 sqthqvslskqptsqsgkeptgpketsgghhhhh 302

RESULT 5  
AAY02350  
ID AAY02350 standard; Protein; 95 AA.  
XX XX  
XX AC AAY02350;

XX 09-JUL-1999 (first entry)  
XX DT  
XX DE A representative HIV-1 Tat-His protein.  
XX KW HIV nef gene; fusion protein; HIV nef gene; Nef protein; Tat protein;  
XX KW vaccine; HIV infection; protein D.  
XX OS Synthetic.  
XX OS Human immunodeficiency virus type 1.

XX PN WO9916884-A1.

XX PD 08-APR-1999.

XX PF 17-SEP-1998; 98WO-EP06040.

XX PR 26-SEP-1997; 97GB-0020585.

XX PA (SMIK ) SMITHKLINE BEECHAM BIOLOGICALS.

XX PI Bruck C, Godart SAG, Marchand M;

XX XX

XX DR WPI; 1999-302282/25.

XX DR N-PSDB; AAX35686.

XX PT HIV Tat or Nef protein linked to a fusion partner

XX PS Disclosure; Fig 2; 66pp; English.

XX CC The present sequence represents a representative HIV-1 Tat-His protein.  
XX CC The protein is used in the creation of the fusion proteins of  
XX CC the invention, in conjunction with a fusion partner (e.g. protein D).  
XX CC The specification also describes fusion proteins comprising HIV-1 Nef  
XX CC protein. The fusion protein can be used in a vaccine to prevent HIV  
XX CC infection.

XX SQ Sequence 95 AA;

Query Match 98.0%; Score 541; DB 20; Length 95;  
Best Local Similarity 96.8%; Pred. No. 4.8e-49;  
Matches 92; Conservative 2; Mismatches 1; Indels 0; Gaps 0;

Qy 1 MEPVDPRLEPWKHPGSOPTACTNCYCKKCCFHCQVCFITAAALGISYGRKKRRRRPPQ 60  
Db 1 mepvdprlepwkhpgsqptactncycckkccfhcqvcfitaalgisyrkrrrrppq 60

Qy 61 GSQTHQVLSKQPTSQSGKEPTGPKETSGHHHHH 95

Db 61 gsqthqvslskqptsqsgkeptgpketsgghhhhh 95

RESULT 6

AAG63232  
ID AAG63232 standard; Protein; 95 AA.

XX AC AAG63232;

XX DT 01-OCT-2001 (first entry)

XX DE Amino acid sequence of a His-tagged Tat protein of HIV.

XX KW HIV; nef gene; tat gene; vaccine; gp120 gene; HIV viral load.

XX OS Synthetic.

XX OS Human immunodeficiency virus.

XX PN WO200154719-A2.

XX PD 02-AUG-2001.

XX PF 29-JAN-2001; 2001WO-EP00944.



XX PR 31-JAN-2000; 2000GB-0002200.  
XX PR 14-APR-2000; 2000GB-0009336.  
XX PR 06-JUN-2000; 2000GB-0013806.  
XX PR 28-JUN-2000; 2000WO-EP05998.  
XX PA (SMIK ) SMITHKLINE BEECHAM BIOLOGICALS.  
XX PI Voss G;  
XX DR WPI: 2001-476172/51.  
XX DR N-PSDB; AAH42876.  
XX PT New use of a human immunodeficiency virus (HIV) Tat, or Nef, or Nef  
XX PT linked to Tat (Nef-Tat) protein or polynucleotide and an HIV gp120  
XX PT protein or polynucleotide for the manufacture of a vaccine -  
XX PS Disclosure; Fig 1; 90pp; English.  
XX CC The present sequence represents a His-tagged Tat protein of HIV. The  
XX CC protein is expressed in the yeast Pichia pastoris, and is used to  
XX CC produce the vaccine of the invention. The specification describes  
XX CC the use of HIV Tat, HIV Nef, or Nef-Tat; and HIV gp120 in the  
XX CC manufacture of a vaccine. The vaccine is used for the prophylactic or  
XX CC therapeutic immunization of humans against HIV. Tat, Nef or Nef-Tat act  
XX CC in synergy with gp120 in the treatment and prevention of HIV. The  
XX CC vaccine reduces the HIV viral load in HIV infected humans and results  
XX CC in a maintenance of CD4+ levels over those levels found in the absence  
XX CC of vaccination with HIV Tat, Nef or Nef-Tat and HIV gp120.  
XX SQ Sequence 95 AA;  
  
Query Match 98.0%; Score 541; DB 22; Length 95;  
Best Local Similarity 96.8%; Pred. No. 4.8e-49;  
Matches 92; Conservative 2; Mismatches 1; Indels 0; Gaps 0;  
  
Qy 1 MEPVDPRLEPKHPSQPKTACTNCYCKKCCFHCQVCFITAAALGYSYGRKKRRRRPPQ 60  
Db 1 mepvdprlepkhpsqpkactncycckkccfhcqvcfitkalgisyrkrrrrppq 60  
  
Qy 61 GSQTHOVSLSKOPTSOSKGEPTGPKETSCHHHHH 95  
Db 61 gsqthovslskoptsqsrgrdtpgpketsghhhhh 95  
  
RESULT 7  
AAY02351  
ID AAY02351 standard; Protein: 302 AA.  
XX AC AAY02351;  
XX DT 09-JUL-1999 (first entry)  
XX DE A representative HIV-1 Nef-Tat-His protein.  
XX KW HIV nef gene; fusion protein; HIV nef gene; Nef protein; Tat protein;  
XX KW vaccine; HIV infection; protein D.  
XX OS Synthetic.  
XX OS Human immunodeficiency virus type 1.  
XX PN WO9916884-A1.  
XX PD 08-APR-1999.  
XX PF 17-SEP-1998; 98WO-EP06040.  
XX PR 26-SEP-1997; 97GB-0020585.  
XX PA (SMIK ) SMITHKLINE BEECHAM BIOLOGICALS.  
XX PI Bruck C, Godart SAG, Marchand M;

XX DR WPI: 1999-302282/25.  
XX DR N-PSDB; AAH35687.  
XX PT HIV Tat or Nef protein linked to a fusion partner  
XX PS Disclosure; Fig 2; 66pp; English.  
XX CC The present sequence represents a representative HIV-1 Nef-Tat-His  
XX CC protein. The protein is used in the creation of the fusion proteins of  
XX CC the invention, in conjunction with a fusion partner (e.g. protein D).  
XX CC The fusion protein can be used in a vaccine to prevent HIV infection.  
XX SQ Sequence 302 AA;  
  
Query Match 97.1%; Score 536; DB 20; Length 302;  
Best Local Similarity 96.8%; Pred. No. 5.3e-48;  
Matches 91; Conservative 2; Mismatches 1; Indels 0; Gaps 0;  
  
Qy 2 EPVDPRLEPKHPSQPKTACTNCYCKKCCFHCQVCFITAAALGYSYGRKKRRRRPPQ 61  
Db 209 epvdprlepkhpsqpkactncycckkccfhcqvcfitkalgisyrkrrrrppq 268  
  
Qy 62 SOTHQVSLSKOPTSOSKGEPTGPKETSCHHHHH 95  
Db 269 sqthqvslskoptsqsrgrdtpgpketsghhhhh 302  
  
RESULT 8  
AAG63233  
ID AAG63233 standard; Protein: 302 AA.  
XX AC AAG63233;  
XX DT 01-OCT-2001 (first entry)  
XX DE Amino acid sequence of a His-tagged Nef-Tat linked protein of HIV.  
XX KW HIV; nef gene; tat gene; vaccine; gp120 gene; HIV viral load.  
XX OS Synthetic.  
XX OS Human immunodeficiency virus.  
XX PN WO200154719-A2.  
XX PD 02-AUG-2001.  
XX PF 29-JAN-2001; 2001WO-EP00944.  
XX PR 31-JAN-2000; 2000GB-0002200.  
XX PR 14-APR-2000; 2000GB-0009336.  
XX PR 06-JUN-2000; 2000GB-0013806.  
XX PR 28-JUN-2000; 2000WO-EP05998.  
XX PA (SMIK ) SMITHKLINE BEECHAM BIOLOGICALS.  
XX PI Voss G;  
XX DR WPI: 2001-476172/51.  
XX DR N-PSDB; AAH42877.  
XX PT New use of a human immunodeficiency virus (HIV) Tat, or Nef, or Nef  
XX PT linked to Tat (Nef-Tat) protein or polynucleotide and an HIV gp120  
XX PT protein or polynucleotide for the manufacture of a vaccine -  
XX PS Disclosure; Fig 1; 90pp; English.  
XX CC The present sequence represents a His-tagged Nef-Tat linked protein of  
XX CC HIV. The protein is expressed in the yeast Pichia pastoris, and is used  
XX CC to produce the vaccine of the invention. The specification describes  
XX CC the use of HIV Tat, HIV Nef, or Nef-Tat; and HIV gp120 in the  
XX CC manufacture of a vaccine. The vaccine is used for the prophylactic or

CC therapeutic immunization of humans against HIV. Tat, Nef or Nef-Tat act  
 CC in synergy with gp120 in the treatment and prevention of HIV. The  
 CC vaccine reduces the HIV viral load in HIV infected humans and results  
 CC in a maintenance of CD4+ levels over those levels found in the absence  
 CC of vaccination with HIV Tat, Nef or Nef-Tat and HIV gp120.  
 XX  
 SQ Sequence 302 AA;

Query Match 97.1%; Score 536; DB 22; Length 302;  
 Best Local Similarity 96.8%; Pred. No. 5.3e-48;  
 Matches 91; Conservative 2; Mismatches 1; Indels 0; Gaps 0;

QY 2 EPVDPRLPEPKHPSGQPKTACTNCYCKCCFHCQVCFTTAALGISYGRKKRRRPPQG 61  
 DB 209 epvdprieplwkhpsgqpkactactncycckccfhcqvcftkalgisygrkkrrirppqg 268  
 QY 62 SOTHQVLSKQPTSQSGKEPTGPKETSGHHHHH 95  
 DB 269 sqthqvsksqptsqsgrdptgpketsghhhhh 302

RESULT 9  
 AAY02353  
 ID AAY02353 standard; Protein; 411 AA.  
 XX  
 AC. AAY02353;  
 XX  
 DT 09-JUL-1999 (first entry)  
 XX  
 DE A representative Lipod-Tat-His fusion protein.  
 XX  
 KW HIV nef gene; fusion protein; HIV nef gene; Nef protein; Tat protein;  
 KW vaccine; HIV infection; protein D.  
 XX  
 OS Synthetic.  
 OS Human immunodeficiency virus type 1.  
 XX  
 PN WO9916884-A1.  
 XX  
 PD 08-APR-1999.  
 XX  
 PF 17-SEP-1998; 98WO-EP06040.  
 XX  
 PR 26-SEP-1997; 97GB-0020585.  
 XX  
 PA (SMIK ) SMITHKLINE BEECHAM BIOLOGICALS.  
 XX  
 PI Bruck C, Godart SAG, Marchand M;  
 XX  
 DR WPI; 1999-302282/25.  
 DR N-PSDB; AAX35689.  
 XX  
 PT HIV Tat or Nef protein linked to a fusion partner  
 XX  
 PS Disclosure; Fig 2; 66pp; English.  
 XX

XX The present sequence represents a fusion protein comprising Lipod-HIV-1  
 CC Tat-His. The protein is exemplified the fusion proteins of  
 CC the invention. The specification also describes fusion proteins  
 CC comprising HIV-1 Nef protein. The fusion protein can be used in a  
 CC vaccine to prevent HIV infection.  
 XX  
 SQ Sequence 411 AA;

Query Match 97.1%; Score 536; DB 20; Length 411;  
 Best Local Similarity 96.8%; Pred. No. 7.3e-48;  
 Matches 91; Conservative 2; Mismatches 1; Indels 0; Gaps 0;

QY 2 EPVDPRLPEPKHPSGQPKTACTNCYCKCCFHCQVCFTTAALGISYGRKKRRRPPQG 61  
 DB 318 epvdprieplwkhpsgqpkactactncycckccfhcqvcftkalgisygrkkrrirppqg 377

QY 62 SOTHQVLSKQPTSQSGKEPTGPKETSGHHHHH 95  
 DB 378 sqthqvsksqptsqsgrdptgpketsghhhhh 411

RESULT 10

AAG63235  
 ID AAG63235 standard; Protein; 411 AA.  
 XX  
 AC. AAG63235;  
 XX  
 DT 01-OCT-2001 (first entry)  
 XX  
 DE Amino acid sequence of a His tagged LipD-Nef-Tat linked protein.  
 XX  
 KW HIV; nef gene; tat gene; vaccine; gp120 gene; HIV viral load.  
 XX  
 OS Synthetic.  
 OS Human immunodeficiency virus.  
 XX  
 FH Key Location/Qualifiers  
 FT Peptide 1..109  
 FT /note= "ProtD fusion partner"  
 XX  
 PN WO200154719-A2.  
 XX  
 PD 02-AUG-2001.  
 XX  
 PF 29-JAN-2001; 2001WO-EP00944.  
 XX  
 PR 31-JAN-2000; 2000GB-0002200.  
 PR 14-APR-2000; 2000GB-0009336.  
 PR 06-JUN-2000; 2000GB-0013806.  
 PR 28-JUN-2000; 2000WO-EP05998.  
 XX  
 PA (SMIK ) SMITHKLINE BEECHAM BIOLOGICALS.  
 XX  
 PI Voss G;  
 XX  
 DR WPI; 2001-476172/51.  
 DR N-PSDB; AAH42879.  
 XX  
 PT New use of a human immunodeficiency virus (HIV) Tat, or Nef, or Nef  
 PT linked to Tat (Nef-Tat) protein or polynucleotide and an HIV gp120  
 PT protein or polynucleotide for the manufacture of a vaccine -  
 XX  
 PS Disclosure; Fig 1; 90pp; English.  
 XX  
 CC The present sequence represents a His-tagged Nef-Tat linked protein of  
 CC HIV, with a lipidation signal sequence (Lipod) which is removed after  
 CC processing and a protD fusion partner. The protein is expressed in  
 CC Escherichia coli, and is used to produce the vaccine of the invention.  
 CC The specification describes the use of HIV Tat, HIV Nef, or Nef-Tat;  
 CC and HIV gp120 in the manufacture of a vaccine. The vaccine is used for  
 CC the prophylactic or therapeutic immunization of humans against HIV.  
 CC Tat, Nef or Nef-Tat act in synergy with gp120 in the treatment and  
 CC prevention of HIV. The vaccine reduces the HIV viral load in HIV  
 CC infected humans and results in a maintenance of CD4+ levels over those  
 CC levels found in the absence of vaccination with HIV Tat, Nef or Nef-Tat  
 CC and HIV gp120.  
 XX  
 SQ Sequence 411 AA;

Query Match 97.1%; Score 536; DB 22; Length 411;  
 Best Local Similarity 96.8%; Pred. No. 7.3e-48;  
 Matches 91; Conservative 2; Mismatches 1; Indels 0; Gaps 0;

QY 2 EPVDPRLPEPKHPSGQPKTACTNCYCKCCFHCQVCFTTAALGISYGRKKRRRPPQG 61  
 DB 318 epvdprieplwkhpsgqpkactactncycckccfhcqvcftkalgisygrkkrrirppqg 377





CC All or part of it can be produced by recombinant vaccinia virus by  
 CC integrating it into the genome region nonessential for vaccinia virus  
 CC proliferation and then infecting animal cells with the recombinant virus.  
 CC Tat protein can be used to diagnose HIV infection and for prophylaxis  
 CC for AIDS or ARC patients. It can also be used to develop a live vaccine  
 CC using temperature-sensitive Lister strain of vaccinia virus.  
 XX  
 SQ Sequence 86 AA;

Query Match 86.6%; Score 478; DB 10; Length 86;

Best Local Similarity 96.5%; Pred. No. 1.6e-42;  
 Matches 83; Conservative 2; Mismatches 1; Indels 0; Gaps 0;

QY 1 MEPVDPRLEPWKHPGSPKACTNCYCKKCCFHCQVCFTAAALGISYGRKKRRQRRPPQ 60  
 DB 1 mepvdprlepwhpgsqptactncycckkccfhcqvcftkaigisygrkkrrrrppq 60  
 QY 61 GSQTHOVSLSKOPTSQSKGPTGPKE 86  
 DB 61 gsqthovslskoptsqsgptgpk 86

#### RESULT 16

AAP91905  
 ID AAP91905 standard; protein; 86 AA.

AC AAP91905;

XX 12-MAY-1990 (first entry)

DE CDNA for tat protein of human immunodeficiency virus (HIV).

XX tat protein; HIV; recombinant vaccinia virus.

XX Human immunodeficiency virus.

FT Key Location/Qualifiers  
 FT CDS 46...306  
 FT /\*tag= a

XX JP01085092-A.

XX 30-MAR-1989.

XX 25-SEP-1987; 87JP-0240513.

XX 25-SEP-1987; 87JP-0240513.

XX (KOKU-) KOKURITSU YOBO EISE (JAPC).

XX WPI; 1989-140777/19.

XX N-PSDB; AAP91905.

XX Prodn. of protein(s)

PT - using recombinant vaccinia viruses contg. protein of

PT Human Immunodeficiency Virus

XX Fig 5; page 863; 9pp; Japanese.

XX It is inserted into the genome of vaccinia viruses in a  
 CC region which is not essential for their propagation to create recombinant  
 CC vaccinia viruses. These viruses are propagated in animal cells together  
 CC with recombinant vaccinia viruses which contain DNA coding for a useful  
 CC protein, for example, human growth hormone, insulin, interferon,  
 CC interleukin-2; haemagglutination factor, hepatitis-B surface antigen,  
 CC amylase, protease. The presence of tat protein assists the mass  
 CC production of the useful protein.

XX Sequence 86 AA;

Query Match 86.6%; Score 478; DB 10; Length 86;

Best Local Similarity 96.5%; Pred. No. 1.6e-42;  
 Matches 83; Conservative 2; Mismatches 1; Indels 0; Gaps 0;

QY 1 MEPVDPRLEPWKHPGSPKACTNCYCKKCCFHCQVCFTAAALGISYGRKKRRQRRPPQ 60  
 DB 1 mepvdprlepwhpgsqptactncycckkccfhcqvcftkaigisygrkkrrrrppq 60

QY 61 GSQTHOVSLSKOPTSQSKGPTGPKE 86

DB 61 gsqthovslskoptsqsgptgpk 86

#### RESULT 17

AAP93140

ID AAP93140 standard; protein; 86 AA.

XX AAP93140;

XX 21-MAR-1990 (first entry)

DE Synthetic TAT protein.

XX TAT protein; HIV-1; tat gene; trans-activator; anti-terminator;

XX WO8909824-A.

XX 19-OCT-1989.

XX 14-APR-1989; 89WO-GB00384.

XX 15-APR-1988; 88GB-0008892.

XX (BRBI-) BRITISH BIOTECHNOLOGY.

XX Edwards RM, Adams SE;

XX WPI; 1989-324229/44.

XX N-PSDB; AAN91837.

PT Synthesis of double stranded DNA, esp. TAT gene of HIV - by prepq. DNA  
 PT contg. a single and double stranded portion and in vivo gap repair.

XX Disclosure; fig.2; 23pp; English.

XX The synthetic tat DNA encoding this protein has useful restriction sites  
 CC to facilitate modification for structural and immunological studies. TAT  
 CC is a transactivator involved in the activation of HIV long terminal  
 CC repeat, acting as an anti-terminator.

XX Sequence 86 AA;

Query Match 86.6%; Score 478; DB 10; Length 86;

Best Local Similarity 96.5%; Pred. No. 1.6e-42;  
 Matches 83; Conservative 2; Mismatches 1; Indels 0; Gaps 0;

QY 1 MEPVDPRLEPWKHPGSPKACTNCYCKKCCFHCQVCFTAAALGISYGRKKRRQRRPPQ 60  
 DB 1 mepvdprlepwhpgsqptactncycckkccfhcqvcftkaigisygrkkrrrrppq 60

QY 61 GSQTHOVSLSKOPTSQSKGPTGPKE 86

DB 61 gsqthovslskoptsqsgptgpk 86

#### RESULT 18

AAR38618

ID AAR38618 standard; protein; 86 AA.

XX AAR38618;

XX 13-DEC-1993 (first entry)

DT 13-DEC-1993 (first entry)

XX

DE Sequence of the HIV-1 TAT protein.  
 XX Human immunodeficiency virus 1; HIV-1; TAT protein; transactivator;  
 KW long terminal repeat; transactivation response element; TAR;  
 KW RNA binding; RNA cleavage.  
 XX  
 XX Human immunodeficiency virus 1.  
 XX  
 FH Key Location/Qualifiers  
 FT Domain 49..57  
 FT /label= nuclear targeting domain  
 FT Region 49..72  
 FT /note= "proteolytic product of wt TAT protein  
 FT which binds specifically to TAR-element-  
 FT contg. RNA"  
 XX  
 PN WO9312234-A.  
 XX  
 XX 24-JUN-1993.  
 PD  
 XX 11-DEC-1992; 92WO-US10770.  
 PF  
 XX 13-DEC-1991; 91US-0808452.  
 PR  
 XX 21-JAN-1992; 92US-0826934.  
 PR  
 XX (STR1 ) SRI INT.  
 PA  
 XX Jayasena SD, Johnston BH;  
 PI  
 XX WPI; 1993-214176/26.  
 DR  
 XX Viral polypeptide(s) with site-specific RNA binding - contain  
 FT motility to cleave RNA backbone and are used to inhibit HIV antigen  
 FT expression in infected cells  
 PT  
 XX Claim 2; Figure 2A; 94pp; English.  
 PS  
 XX TAT is a potent transactivator of long terminal repeat (LTR)-  
 XX directed viral gene expression. TAT-induced transactivation requires  
 CC the present of the TAR (transactivation response) element, located  
 CC at the 5'UTR of the viral mRNA element. The sequence of RNA target  
 CC molecules were chosen based on previous studies characterising the  
 CC binding properties of the HIV-encoded TAT protein and the TAR target  
 CC region from both HIV-1 and HIV-2. HIV-1 TAR is the 57-nt RNA stem-  
 CC loop structure found in HIV-1 mRNA (nt 1-57). The RNA substrate is  
 CC given in AAQ44139.  
 XX  
 SQ Sequence 86 AA;

Query Match 86.6%; Score 478; DB 14; Length 86;  
 Best Local Similarity 96.5%; Pred. No. 1.6e-42;  
 Matches 83; Conservative 2; Mismatches 1; Indels 0; Gaps 0;

OY 1 MEPVDPRLEPKHPSQPKTACTNCYCKKCCFHCQVCFITAAALGISYGRKKRRRRPPQ 60  
 DB 1 mepvdprlepkhpsqpkactncycckccfhcqvcfitaalgisyrkrrrrppq 60  
 OY 61 GSQTHQVSLSKOPTSQSGEPTGPKE 86  
 DB 61 gsqthqvskskptsqsgdptgpk 86

RESULT 19  
 AAR86625  
 ID AAR86625 standard; Protein; 86 AA.  
 XX  
 AC AAR86625;  
 XX  
 DT 28-JUN-1996 (first entry)  
 XX  
 DE HIV TAT.  
 XX

KW Immunogen; cellular uptake region; transactivating protein; TAT; HIV;  
 KW viraemia; antibody.  
 XX  
 OS Human immunodeficiency virus.  
 XX  
 FH Key Location/Qualifiers  
 FT Peptide 46..63  
 FT /note= "Cellular uptake region"  
 XX  
 XX WO9531999-A1.  
 PN  
 XX 30-NOV-1995.  
 PD  
 XX 16-MAY-1995; 95WO-US06077.  
 PF  
 XX 23-MAY-1994; 94US-0247991.  
 PR  
 XX (IMMU-) IMMUNOGIOLOGY RES INST INC.  
 PA  
 XX Culler MD, Goldstein G, Shenbagamurthi P;  
 PI  
 XX WPI; 1996-049298/05.  
 DR  
 XX N-PSDB; AAT06634.  
 DR  
 XX Immunogen containing cellular uptake region of viral TAT protein -  
 PT induces high antibody titre against TAT, partic. for preventing or  
 PT treating HIV infection  
 XX  
 PS Disclosure; Figure 1; 54pp; English.  
 PS  
 XX This sequence represents the HIV transactivating protein, TAT. The  
 CC sequences given in AAR86601-19 are immunogens derived from the cellular  
 CC uptake region of this protein. These peptides are based on the region  
 CC comprising amino acids 46-63. The immunogenic peptides are used to  
 CC protect against infection by HIV. They may also be used to reduce  
 CC viraemia in already infected patients. These immunogens may be used  
 CC to identify, or generate, antibodies specific for TAT.  
 XX  
 SQ Sequence 86 AA;

Query Match 86.6%; Score 478; DB 17; Length 86;  
 Best Local Similarity 96.5%; Pred. No. 1.6e-42;  
 Matches 83; Conservative 2; Mismatches 1; Indels 0; Gaps 0;

OY 1 MEPVDPRLEPKHPSQPKTACTNCYCKKCCFHCQVCFITAAALGISYGRKKRRRRPPQ 60  
 DB 1 mepvdprlepkhpsqpkactncycckccfhcqvcfitaalgisyrkrrrrppq 60  
 OY 61 GSQTHQVSLSKOPTSQSGEPTGPKE 86  
 DB 61 gsqthqvskskptsqsgdptgpk 86

RESULT 20  
 AAW31202  
 ID AAW31202 standard; protein; 86 AA.  
 XX  
 AC AAW31202;  
 XX  
 DT 20-MAR-1998 (first entry)  
 XX  
 DE Human Immunodeficiency Virus Type 1 Tat protein.  
 XX  
 KW Human immunodeficiency virus; HIV Type 1; Tat protein;  
 KW cargo molecules; intracellular delivery; fusion protein;  
 KW therapeutic; prophylactic; diagnostic; transport polypeptide.  
 XX  
 OS Human immunodeficiency virus type 1.  
 XX  
 FH Key Location/Qualifiers  
 FT Region 22..36  
 FT /label= cysteine-rich

FT Region 49..57  
 FT Domain /label= basic\_region  
 FT 73..86  
 FT /label= carboxy\_terminal\_domain

PN US5674980-A.  
 XX  
 XX  
 PD 07-OCT-1997.

PF 21-DEC-1989; 89US-0454450.

PR 28-APR-1994; 94US-0235403.  
 PR 21-DEC-1989; 89US-0454450.  
 PR 02-JAN-1991; 91US-0636662.  
 PR 21-AUG-1992; 92US-0934375.  
 PR 19-AUG-1993; 93WO-US07833.  
 PR 24-NOV-1993; 93US-0158015.  
 PR 25-MAY-1995; 95US-0450098.

XX (BARS/) BARSOUM J G.  
 PA (FAWE/) FAWELL S E.  
 PA (FRAN/) FRANKEL A.  
 PA (PABO/) PABO.  
 PA (PEPI/) PEPINSKY R B.

PI Barsoum JG, Fawell SE, Frankel A, Pabo C, Pepinsky RB;

XX WPI; 1997-502388/46.

XX Fusion proteins containing truncated HIV tat sequences - useful for  
 PT intracellular delivery of viral repressor proteins

XX Example I; Fig 1; 77pp; English.

XX This sequence represents the Tat protein from Human Immunodeficiency  
 CC virus, Type 1 which can be used in a novel method for the delivery of  
 CC biologically active cargo molecules into the cytoplasm and nuclei of  
 CC eukaryotic cells. The tat protein from immunodeficiency virus (e.g.  
 CC HIV-1, HIV-2, SIV) is readily taken up into cells when present  
 CC extracellularly. By modifying the HIV tat protein i.e. removing the  
 CC cysteine-rich region and the carboxy-terminal domain and covalently  
 CC linking the protein to a cargo molecule to form a fusion protein,  
 CC the problems of spurious trans-activation and disulphide aggregation are  
 CC eradicated. The reduced size of transport polypeptides also minimises  
 CC interference with the biological activity of the cargo molecule. This is  
 CC applicable for therapeutic, prophylactic or diagnostic intracellular  
 CC delivery of small molecules and macromolecules e.g. proteins, nucleic  
 CC acids and polysaccharides.

XX Sequence 86 AA;

Query Match 86.6%; Score 478; DB 18; Length 86;  
 Best Local Similarity 96.5%; Pred. No. 1.6e-42;  
 Matches 83; Conservative 2; Mismatches 1; Indels 0; Gaps 0;

QY 1 MEVPDPRLPEPKHPSQPKTACTNCYCKKCFHCQVCFITAAIGISYGRKKRRRPPQ 60  
 Db |||||

QY 61 GSQTHQVLSLKOPTSQSKGEPTGPK 86

Db 61 gsqthqvskskptsqsrqdgptgpk 86

RESULT 21  
 ID AAW26443  
 XX AAW26443 standard; Protein; 86 AA.

AC AAW26443;

XX 16-DEC-1997 (first entry)

DE HIV tat protein.

XX HIV; tat protein; transport protein; cargo delivery.

OS Human immunodeficiency virus type 1.

XX Key Location/Qualifiers  
 FT Region 23..36  
 FT /label= Cys-rich\_region  
 FT Region 47..59  
 FT /label= Basic\_region  
 FT Region 73..86  
 FT /label= C-terminal

XX US5652122-A.

PN 29-JUL-1997.

PD 21-DEC-1989; 89US-0454450.

PR 28-APR-1994; 94US-0235403.  
 PR 21-DEC-1989; 89US-0454450.  
 PR 02-JAN-1991; 91US-0636662.  
 PR 21-AUG-1992; 92US-0934375.  
 PR 19-AUG-1993; 93WO-US07833.  
 PR 24-NOV-1993; 93US-0158015.  
 PR 25-MAY-1995; 95US-0450257.

XX (BARS/) BARSOUM J G.

PA (FAWE/) FAWELL S E.

PA (FRAN/) FRANKEL A.

PA (PABO/) PABO C.

PA (PEPI/) PEPINSKY R B.

XX Barsoum JG, Fawell SE, Frankel A, Pabo C, Pepinsky RB;

XX WPI; 1997-392943/36.

XX New DNA constructs for transporting molecules to cells - encode a  
 PT fusion protein comprising a modified HIV tat protein and a  
 PT carboxy-terminal cargo moiety

XX Disclosure; Fig 1; 76pp; English.

XX This polypeptide comprises the HIV tat protein. Novel transport  
 CC proteins comprise modified HIV tat covalently attached to a cargo  
 CC molecule (see AAW26436-42). The modified tat protein (see AAW26444-49)  
 CC is characterised by the presence of the tat basic region (thought  
 CC to be required for nuclear localisation), and the absence of the  
 CC Cys-rich region and the exon 2-encoded C-terminal region. These  
 CC modifications solve the potential problems of spurious  
 CC trans-activation and disulphide aggregation, while the reduced size  
 CC of the transport proteins minimises interference with the  
 CC biological activity of the cargo molecule. DNA molecules that  
 CC encode the modified tat fusion proteins are claimed and can be  
 CC used to deliver polypeptides or nucleic acids to the cytoplasm  
 CC of nuclei of cells in vivo or in vitro.

XX Sequence 86 AA;

Query Match 86.6%; Score 478; DB 18; Length 86;  
 Best Local Similarity 96.5%; Pred. No. 1.6e-42;  
 Matches 83; Conservative 2; Mismatches 1; Indels 0; Gaps 0;

QY 1 MEVPDPRLPEPKHPSQPKTACTNCYCKKCFHCQVCFITAAIGISYGRKKRRRPPQ 60  
 Db |||||

QY 61 GSQTHQVLSLKOPTSQSKGEPTGPK 86

Db 61 gsqthqvskskptsqsrqdgptgpk 86

## RESULT 22

AAW71385  
ID AAW71385 standard; Protein: 86 AA.XX  
AC AAW71385;XX  
DT 04-DEC-1998 (first entry)XX  
DE Amino acid sequence of HIV Tat 85.XX  
KW HIV Tat 85; leaderless protein; inhibition; export; angiogenesis;  
KW restenosis; treatment; tumour; inflammation; cell proliferation;  
KW diabetes; retinopathy; infection; polycystic kidney disease;  
KW atherosclerosis.XX  
OS Human immunodeficiency virus.XX  
FN WO9837880-A1.XX  
PD 03-SEP-1998.XX  
PF 25-FEB-1998; 98WO-US03689.XX  
PR 26-FEB-1997; 97US-0807014.XX  
PA (CIBL-) CIBLEX CORP.XX  
PI Baird A, Florkiewicz RZ;XX  
DR WPI; 1998-495377/42.XX  
DR N-PSDB; AAV60345.XX  
PT Inhibiting export of leaderless protein with agent that inhibits  
PT binding to transporter protein - especially for treating  
PT angiogenesis and restenosis by preventing export of fibroblast  
PT growth factor, also methods for identifying leaderless proteins and  
PT their transportersXX  
PS Claim 2; Page 70; 116pp; English.XX  
CC The present sequence represents Human Immunodeficiency virus (HIV) Tat  
CC 85, a leaderless protein. A leaderless protein refers to a protein that  
CC is found in an extracellular environment, but lacks a canonical leader  
CC sequence. The specification describes a method for inhibiting export of  
CC a leaderless protein from a cell. The method comprises treating  
CC the cell with an agent that inhibits binding between the leaderless  
CC protein and a transport molecule. Treatment with the inhibiting agent  
CC is specifically used to treat angiogenesis and restenosis, i.e. where  
CC expression of FGF-2 is inhibited, and the agent is applied to endothelial  
CC or smooth muscle cells. Other applications are treatment of tumours  
CC (melanoma, teratocarcinoma, ovarian carcinoma, bladder cancer and  
CC neuroblastoma), inflammation, cell proliferation, complications of  
CC diabetes (e.g. retinopathy), viral, bacterial or fungal infections,  
CC polycystic kidney disease and atherosclerosis.XX  
SQ Sequence 86 AA;Query Match 86.6%; Score 478; DB 19; Length 86;  
Best Local Similarity 96.5%; Pred. No. 1.6e-42;  
Matches 83; Conservative 2; Mismatches 1; Indels 0; Gaps 0;

Qy 1 MEPVDPRLEPKWPGSQPKTACTNCYCKKCCFHCQVCFITTAALGTSYGRKKRRRRPPQ 60

Db 1 mepvdprlepkwpgsqpktactncycckccfchcqvcfitkalgisyrkrrrrppq 60  
|||||

Qy 61 GSQTHQVSLSKQPTSQSKGEGTPKME 86

Db 61 gsqthqvskskptsqsgtgdtpgkpe 86  
|||||

## RESULT 23

AAW76148

ID AAW76148 standard; protein: 86 AA.

XX  
AC AAW76148;XX  
DT 24-NOV-1998 (first entry)XX  
DE HIV Type I TAT protein.XX  
KW TAT protein; cargo molecule; therapy; diagnosis; transport protein;  
KW fusion protein; human papillomavirus E2 repressor; target cell.XX  
OS Human immunodeficiency virus type 1.XX  
FN US5804604-A.XX  
PD 08-SEP-1998.XX  
PF 25-MAY-1995; 95US-0450236.XX  
PR 28-APR-1994; 94US-0235403.XX  
PR 21-DEC-1989; 89US-0454450.XX  
PR 02-JAN-1991; 91US-0636662.XX  
PR 19-AUG-1993; 93WO-US07833.XX  
PR 24-NOV-1993; 93US-0158015.XX  
PR 25-MAY-1995; 95US-0450236.XX  
PA (BIOJ ) BIOGEN INC.XX  
PI Barsoum JG, Fawell SE, Frankel A, Pabo C, Pepinsky RB;XX  
DR WPI; 1998-505702/43.XX  
PT HIV tat-derived transport fusion proteins - used to deliver  
PT biological active molecules e.g. peptide(s) or nucleic acids,  
PT specifically into cytoplasm or nuclei of cellsXX  
PS Disclosure: Fig 1; 83pp; English.XX  
CC This sequence represents the human immunodeficiency virus (HIV) Type I  
CC TAT protein which is used in a method for the delivery of biologically  
CC active cargo molecules into the cytoplasm and nuclei of cells, for  
CC therapeutic, prophylactic or diagnostic purposes. This is accomplished  
CC by the presence of a small, basic section of tat transport protein of  
CC HIV. This is used as it is this protein which is observed to cause human  
CC cells in culture to take up HIV. The method involves the use of a cargo  
CC moiety in combination with a transport moiety usually in the form of a  
CC fusion protein. The cargo moiety is a human papillomavirus E2 repressor  
CC and where the transport moiety is one of following HIV tat protein  
CC fragments (a) aa 47-58, (b) aa 47-72, (c) 38-72, (d) aa 38-58, (e)  
CC aa 37-58, (f) aa 1-21 and 38-72, (g) aa 47-62 or aa 38-62. The proteins  
CC allow delivery of specific peptides into cells at high concentrations  
CC due to use of existing transporters. Previous methods of delivery  
CC include bombardment and transforming, which only allow a fraction of the  
CC cell population to be infected and can additionally damage cells as  
CC they cause physical opening of the cell walls/membranes to allow entry.XX  
SQ Sequence 86 AA;Query Match 86.6%; Score 478; DB 19; Length 86;  
Best Local Similarity 96.5%; Pred. No. 1.6e-42;  
Matches 83; Conservative 2; Mismatches 1; Indels 0; Gaps 0;

Qy 1 MEPVDPRLEPKWPGSQPKTACTNCYCKKCCFHCQVCFITTAALGTSYGRKKRRRRPPQ 60

Db 1 mepvdprlepkwpgsqpktactncycckccfchcqvcfitkalgisyrkrrrrppq 60  
|||||

Qy 61 GSQTHQVSLSKQPTSQSKGEGTPKME 86

Db 61 gsqthqvskskptsqsgtgdtpgkpe 86  
|||||



```
RESULT 24
ID AAY96881 standard; Protein: 86 AA.
AC AAY96881;
XX
XX
XX 31-OCT-2000 (first entry)
XX
XX Human immunodeficiency virus Tat 85.
XX
XX Human immunodeficiency virus; HIV; Tat 85; transport molecule; golgi;
XX leaderless; endoplasmic reticulum; protein export; detection; inhibitor.
XX
XX Human immunodeficiency virus type 1.
XX
XX Key Location/Qualifiers
XX Misc-difference 53
XX FT /note= "Encoded by AGAA"
XX FT Misc-difference 80
XX FT /note= "Encoded by GAA"
XX
XX USG083706-A.
XX
XX 04-JUL-2000.
XX
XX 25-FEB-1998; 98US-0030613.
XX
XX 26-FEB-1997; 97US-0807014.
XX
XX (CIBL-) CIBLEX CORP.
XX
XX Baird A, Florkiewicz RZ;
XX
XX WPI; 2000-464338/40.
XX
XX N-PSDB; AAA53572.
XX
XX Detecting transport molecules, useful for identifying proteins that
XX mediate leaderless protein export across cell membranes, by contacting
XX cell extracts with a fusion protein of leaderless protein and a tag to
XX form a complex
XX
XX Disclosure; Column 55-56; 64pp; English.
XX
XX Detecting a transport molecule involved in non-endoplasmic reticulum
XX (ER)/Golgi leaderless protein export, comprises contacting test cell
XX extracts or membranes with a fusion protein of a leaderless protein and
XX a tag to form a complex of the fusion protein bound to the transport
XX molecule, and detecting the transport molecule in an isolated complex.
XX The leaderless protein is a protein found in the extracellular
XX environment that lacks a canonical leader sequence, interleukin (IL)
XX 1-alpha, or 1-beta, fibroblast growth factor (FGF) 1 or 2, human
XX immunodeficiency virus (HIV) tat, platelet-derived endothelial cell
XX growth factor (PD-ECGF), ciliary neurotrophic factor (CNTF), sciatic
XX nerve growth-promoting activity, vas deferens protein, transglutaminase,
XX L-14 lectin, factor XIIIa, thiodoxin-like protein (ADF), thymosin,
XX parathyromosin, mammary-derived growth inhibitor, galectin or rhodanase.
XX The method is used to detect proteins, complexes of proteins, or parts of
XX a larger complex, that bind to and mediate the transport of leaderless
XX proteins, e.g. Na+/K+ ATPase which is an integral membrane protein
XX responsible for transporting sodium and potassium ions across the cell
XX membrane using ATP as the driving force. Transport molecules detected by
XX the method are used in assays to identify inhibitors of the interaction
XX with a leaderless protein.
XX
XX Sequence 86 AA:
XX
XX Query Match 86.6%; Score 478; DB 21; Length 86;
XX Best Local Similarity 96.5%; Pred. No. 1.6e-42;
XX Matches 83; Conservative 2; Mismatches 1; Indels 0; Gaps 0;
XX
XX 1 MEPVDPRLEPWKHPGSQLPACTNCKKCCFHCQVCFTALGTSYGRKKRRRPPQ 60
```

```
Db 1 mepvdprlepwkhpqsgpqtactnckkccfhcqvcftalgsygrkkrrrrppq 60
Qy 61 GSQTHQVLSKQPTSQSKGPTGPK 86
Db 61 gsqthqvskskqptsgsgdptgpk 86

RESULT 25
AAB85996
ID AAB85996 standard; Protein: 86 AA.
XX
XX AAB85996;
XX
XX 30-NOV-2001 (first entry)
XX
XX Amino acid sequence of HIV-1 isolate BH10 tat protein.
XX
XX HIV-1; gpl20; BH10; vaccine; immunization; tat protein.
XX
XX Human immunodeficiency virus type 1.
XX
XX USG268484-B1.
XX
XX 31-JUL-2001.
XX
XX 30-JUL-1998; 98US-0124900.
XX
XX 07-JUN-1995; 95US-0478536.
XX
XX 19-APR-1995; 95WO-EP01481.
XX
XX (POLY-) POLYMUN SCI IMMUNOBIOLOGISCHE FORSCHUNG.
XX
XX Katinger H, Buchacher A, Ernst W, Ballaun C, Purtscher M;
XX Trkola A, Predl R, Schmatz C, Klima A, Steindl F, Muster T;
XX
XX WPI; 2001-556601/62.
XX
XX N-PSDB; AAH76385.
XX
XX New anti-idiotypic antibodies consisting of one or both amino acid
XX sequences corresponding to amino acid positions 79-84 or 326-400 of the
XX processed gp120 of HIV-isolate BH10, useful as vaccine against HIV-1
XX infections
XX
XX Claim 1; Columns 31-32; 27pp; English.
XX
XX The invention relates to a peptide fragment comprising of amino acid
XX sequences corresponding to sequences within the processed gp120 of HIV-1
XX isolate BH10 (GenBank accession M15654). The peptides are useful in the
XX detection, prevention and treatment of HIV-1 infections, and in AIDS
XX therapy. The antibodies are especially useful as vaccines for active and
XX passive immunization, or for the detection and/or determination of HIV-1
XX infected cells and/or HIV-1 viruses. The present sequence represents the
XX amino acid sequence of a tat protein from HIV-1 isolate BH10
XX (GenBank accession M15654).
XX
XX Sequence 86 AA:
XX
XX Query Match 86.6%; Score 478; DB 22; Length 86;
XX Best Local Similarity 96.5%; Pred. No. 1.6e-42;
XX Matches 83; Conservative 2; Mismatches 1; Indels 0; Gaps 0;
XX
XX 1 MEPVDPRLEPWKHPGSQLPACTNCKKCCFHCQVCFTALGTSYGRKKRRRPPQ 60
Db 1 mepvdprlepwkhpqsgpqtactnckkccfhcqvcftalgsygrkkrrrrppq 60
Qy 61 GSQTHQVLSKQPTSQSKGPTGPK 86
Db 61 gsqthqvskskqptsgsgdptgpk 86

RESULT 26
```

```

AARL13379 standard; Protein; 87 AA.
XX
XX AARL13379;
XX
XX DT 23-OCT-1991 (first entry)
XX
XX DE TAT protein.
XX
XX KW HIV; AIDS.
XX
XX OS Synthetic.
XX
XX PN EP441582-A.
XX
XX PD 14-AUG-1991.
XX
XX PF 04-FEB-1991; 91EP-0300903.
XX
XX PR 09-FEB-1990; 90GB-0003010.
XX
XX PA (GLAX ) GLAXO GROUP LTD.
XX
XX PI Dykes CW, Ernst JF, Hobden AN;
XX
XX DR WPI; 1991-240122/33.
XX
XX DR N-PSDB; AAQ13188.
XX
XX PT Gene expression system for yeast cells - with TAT protein
XX mediated expression under the control of a HIV regulatory control
XX sequence.
XX
XX PS Disclosure; Page 10; 26pp; English.
XX
XX CC The TAT protein here is encoded by the synthetic tat gene but is
XX identical to the Tat protein found in HIV. The expresion of the Tat
XX protein in transformed yeast enables Tat inhibitory agents to be
XX screened to obtain potential anti-viral agents esp. anti-HIV
XX therapeutic agents.
XX
XX SQ Sequence 87 AA;

Query Match      86.6%; Score 478; DB 12; Length 87;
Best Local Similarity 96.5%; Pred. No. 1.6e-42;
Matches 83; Conservative 2; Mismatches 1; Indels 0; Gaps 0;

Qy 1 MEPVDPRLEPWKPGSPKTAICTNCKYKKCFHCQVCFTAAALGISYGRKKRRRPPQ 60
   | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
Db 1 mepvdprlepwkpgspktaictnckykccfqcfcitkalgisyrkkrrrppq 60
   | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |

Qy 61 GSOTHQVSLSKQPTSQSKEGTPGKE 86
   | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
Db 61 gsqthqvslskqptsqsrgdptgpk 86
   | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |

RESULT 27
RAY22959
ID RAY22959 standard; Protein; 86 AA.
XX
XX AC RAY22959;
XX
XX DT 20-AUG-1999 (first entry)
XX
XX DE Cys22 mutant tat amino acid sequence of HIV-1.
XX
XX KW HIV Tat protein; AIDS; tumour; HIV infection; dendritic cell;
KW Kaposi's sarcoma cell; activated endothelial cell;
KW cytokine-activated endothelial cell; vaccine.
XX
XX OS Human immunodeficiency virus type 1.
XX
XX PN WO9927958-A2.
```



CC syndromes and symptoms associated with HIV infection. A biologically  
CC active Tat protein, fragments and/or mutants and/or Tat DNA which is  
CC capable of entering and localizing in the nuclei of activated endothelial  
CC cells or dendritic cells and/or activating the proliferation, migration  
CC and invasion of Kaposi's sarcoma (KS) cells and cytokine-activated  
CC endothelial cells, is used as a vaccine. HIV-1 Tat or its mutants  
CC in a biologically active form is able to induce a very strong immune  
CC response against HIV, able to prevent infection or the development of  
CC the disease and to permit efficient therapeutic strategies in  
CC HIV-1-infected individuals.

XX  
SQ Sequence 83' AA;

Query Match 82.9%; Score 457.5; DB 20; Length 83;  
Best Local Similarity 95.3%; Pred. No. 2.le-40;  
Matches 82; Conservative 0; Mismatches 1; Indels 3; Gaps 1;

Qy 1 MEPVDPRLEPWKHPGSPQKTACTNCTCYKCCFHCQVCFITAAALGISYGRKKRRRRPPQ 60  
Db 1 mepvdprlepwkhpqsgqktactncycckccfhcqvcfittalgisysgrkrrrrppq 60  
Qy 61 GSQTHQVSLSKOPTSQSGKEPTGPKE 86  
Db 61 gsqthqvslskqptsqs---ptgpke 83

Search completed: August 26, 2002, 08:12:31  
Job time: 240 sec

GenCore version 4.5  
Copyright (c) 1993 - 2000 Compugen Ltd.

OM protein - protein search, using sw model

Run on: August 26, 2002, 08:10:41 ; Search time 38.46 Seconds  
(without alignments)  
60.334 Million cell updates/sec

Title: US-09-509-239-23  
Perfect score: 552  
Sequence: 1 MEPVDRLEPWKHPGSPKPT.....QSKGPTGPKETSGHHHHH 95

Scoring table: BLOSUM62  
Gapop 10.0 , Gapext 0.5

Searched: 231628 seqs, 24425594 residues

Total number of hits satisfying chosen parameters: 231628

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

#### Database :

Issued\_Patents\_AA:\*  
1: /cgn2\_6/ptodata/2/iaa/5A\_COMB.pep.\*  
2: /cgn2\_6/ptodata/2/iaa/5B\_COMB.pep.\*  
3: /cgn2\_6/ptodata/2/iaa/6A\_COMB.pep.\*  
4: /cgn2\_6/ptodata/2/iaa/6B\_COMB.pep.\*  
5: /cgn2\_6/ptodata/2/iaa/PCTUS.COMB.pep.\*  
6: /cgn2\_6/ptodata/2/iaa/backfiles1.pep.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

#### SUMMARIES

Result No.	Score	Query Match %	Length	ID	Description
1	478	86.6	86	1	US-08-450-257-1
2	478	86.6	86	1	US-08-450-246-1
3	478	86.6	86	1	US-08-450-098-1
4	478	86.6	86	1	US-08-451-233-1
5	478	86.6	86	1	US-08-450-236-1
6	478	86.6	86	3	US-07-808-452-1
7	478	86.6	86	3	US-09-030-613-19
8	478	86.6	86	4	US-09-124-900-6
9	478	86.6	86	4	US-09-451-905-19
10	478	86.6	86	4	US-08-235-403-1
11	478	86.6	86	5	PCT-US92-10770-1
12	478	86.6	86	5	PCT-US92-06077-2
13	462	83.7	86	2	US-08-505-210-1
14	462	83.7	86	4	US-09-089-333-1
15	447	81.0	82	1	US-08-053-079A-15
16	408	73.9	72	3	US-09-030-613-17
17	408	73.9	72	4	US-09-451-905-17
18	384	69.6	72	2	US-08-893-853-1
19	384	69.6	72	4	US-09-113-921-1
20	360.5	65.3	312	1	US-08-094-128A-27
21	360.5	65.3	312	1	US-08-455-674-27
22	360.5	65.3	312	1	US-08-455-992-27
23	360.5	65.3	312	1	US-08-455-972-27
24	360.5	65.3	312	5	PCT-US92-00652-27
25	351	63.6	72	1	US-07-910-867B-2
26	330	59.8	72	1	US-07-910-867B-5
27	324	58.7	72	1	US-07-910-867B-3

28	310	56.2	72	1	US-07-910-867B-1	Sequence 1, Appl
29	276	50.0	56	1	US-08-450-257-7	Sequence 7, Appl
30	276	50.0	56	1	US-08-450-246-7	Sequence 7, Appl
31	276	50.0	56	1	US-08-450-098-7	Sequence 7, Appl
32	276	50.0	56	1	US-08-451-233-7	Sequence 7, Appl
33	276	50.0	56	1	US-08-450-236-7	Sequence 7, Appl
34	276	50.0	56	4	US-08-235-403-7	Sequence 7, Appl
35	188	34.1	37	1	US-07-724-500B-22	Sequence 22, Appl
36	181	32.8	36	1	US-08-450-257-2	Sequence 2, Appl
37	181	32.8	36	1	US-08-450-246-2	Sequence 2, Appl
38	181	32.8	36	1	US-08-450-098-2	Sequence 2, Appl
39	181	32.8	36	1	US-08-451-233-2	Sequence 2, Appl
40	181	32.8	36	1	US-08-450-236-2	Sequence 2, Appl
41	181	32.8	36	4	US-08-235-403-2	Sequence 2, Appl
42	181	32.8	36	4	US-09-336-093-6	Sequence 6, Appl
43	180.5	32.7	157	1	US-08-450-257-60	Sequence 60, Appl
44	180.5	32.7	157	1	US-08-450-246-60	Sequence 60, Appl
45	180.5	32.7	157	1	US-08-450-098-60	Sequence 60, Appl

#### ALIGNMENTS

RESULT 1  
US-08-450-257-1  
; Sequence 1, Application US/08450257  
; Patent No. 5652122  
; GENERAL INFORMATION:  
; APPLICANT: FRANKEL, Alan  
; APPLICANT: PABO, Carl  
; APPLICANT: BARSOUM, James G.  
; APPLICANT: FAWELL, Stephen E.  
; APPLICANT: PEPINSKY, R. B.  
; TITLE OF INVENTION: TAT-DERIVED TRANSPORT POLYPEPTIDES  
; NUMBER OF SEQUENCES: 69  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: FISH & NEAVE  
; STREET: 1251 Avenue of the Americas  
; CITY: New York  
; STATE: New York  
; COUNTRY: USA  
; ZIP: 10020  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Floppy disk  
; COMPUTER: IBM PC compatible  
; OPERATING SYSTEM: PC-DOS/MS-DOS  
; SOFTWARE: PatentIn Release #1.0, Version #1.25  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/08/450.257  
; FILING DATE: 25-MAY-1995  
; CLASSIFICATION: 514  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: US 08/235,403  
; FILING DATE: 28-APR-1994  
; APPLICATION NUMBER: US 07/934,375  
; FILING DATE: 21-AUG-1992  
; APPLICATION NUMBER: US 07/098,766  
; FILING DATE: 28-JUL-1993  
; APPLICATION NUMBER: PCT/US93/07833  
; FILING DATE: 19-AUG-1993  
; APPLICATION NUMBER: US 07/454,450  
; FILING DATE: 21-DEC-1989  
; APPLICATION NUMBER: US 07/636,662  
; FILING DATE: 02-JAN-1991  
; APPLICATION NUMBER: US 08/158,015  
; FILING DATE: 24-NOV-1993  
; ATTORNEY/AGENT INFORMATION:  
; NAME: Haley Jr., James F.  
; REGISTRATION NUMBER: 27,794  
; REFERENCE/DOCKET NUMBER: B170 CIP 2  
; TELECOMMUNICATION INFORMATION:  
; TELEPHONE: (212) 596-9000  
; TELEFAX: (212) 596-9090

```
TELEX: 14-8367
INFORMATION FOR SEQ ID NO: 1:
SEQUENCE CHARACTERISTICS:
LENGTH: 86 amino acids
TYPE: amino acid
TOPOLOGY: linear
MOLECULE TYPE: protein
ORIGINAL SOURCE:
ORGANISM: human immunodeficiency virus
STRAIN: type 1
US-08-450-257-1

Query Match      86.6%; Score 478; DB 1; Length 86;
Best Local Similarity 96.5%; Pred. No. 5.4e-46;
Matches 83; Conservative 2; Mismatches 1; Indels 0; Gaps 0;

Qy 1 MEPVDPRLEPKHGPSQPKTACTNCYCKKCCFHCQVCFITAAALGISYGRKKRRRRPPQ 60
Db 1 MEPVDPRLEPKHGPSQPKTACTNCYCKKCCFHCQVCFITAAALGISYGRKKRRRRPPQ 60
Qy 61 GSQTHQVSLSKOPTSQSKGPTGPKE 86
Db 61 GSQTHQVSLSKOPTSQSKGPTGPKE 86

RESULT 2
US-08-450-246-1
Sequence 1, Application US/08450246
Patent No. 5670617
GENERAL INFORMATION:
APPLICANT: FRANKEL, Alan
APPLICANT: PABO, Carl
APPLICANT: BARSOUM, James G.
APPLICANT: FAWELL, Stephen E.
APPLICANT: PEPINSKY, R. B.
TITLE OF INVENTION: TAT-DERIVED TRANSPORT POLYPEPTIDES
NUMBER OF SEQUENCES: 69
CORRESPONDENCE ADDRESS:
ADDRESSEE: FISH & NEAVE
STREET: 1251 Avenue of the Americas
CITY: New York
STATE: New York
COUNTRY: USA
ZIP: 10020
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patentin Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/450.246
FILING DATE: 25-MAY-1995
CLASSIFICATION: 514
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/235.403
FILING DATE: 28-APR-1994
APPLICATION NUMBER: US 07/934.375
FILING DATE: 21-AUG-1992
APPLICATION NUMBER: US 07/098.766
FILING DATE: 28-JUL-1993
APPLICATION NUMBER: PCT/US93/07833
FILING DATE: 19-AUG-1993
APPLICATION NUMBER: US 07/454.450
FILING DATE: 21-DEC-1989
APPLICATION NUMBER: US 07/636.662
FILING DATE: 02-JAN-1991
APPLICATION NUMBER: US 08/158.015
FILING DATE: 24-NOV-1993
ATTORNEY/AGENT INFORMATION:
NAME: Haley Jr., James F.
REGISTRATION NUMBER: 27,794
REFERENCE/DOCKET NUMBER: B170 CIP 2
```

```
TELECOMMUNICATION INFORMATION:
TELEPHONE: (212) 596-9000
TELEFAX: (212) 596-9090
TELEX: 14-8367
INFORMATION FOR SEQ ID NO: 1:
SEQUENCE CHARACTERISTICS:
LENGTH: 86 amino acids
TYPE: amino acid
TOPOLOGY: linear
MOLECULE TYPE: protein
ORIGINAL SOURCE:
ORGANISM: human immunodeficiency virus
STRAIN: type 1
US-08-450-246-1

Query Match      86.6%; Score 478; DB 1; Length 86;
Best Local Similarity 96.5%; Pred. No. 5.4e-46;
Matches 83; Conservative 2; Mismatches 1; Indels 0; Gaps 0;

Qy 1 MEPVDPRLEPKHGPSQPKTACTNCYCKKCCFHCQVCFITAAALGISYGRKKRRRRPPQ 60
Db 1 MEPVDPRLEPKHGPSQPKTACTNCYCKKCCFHCQVCFITAAALGISYGRKKRRRRPPQ 60
Qy 61 GSQTHQVSLSKOPTSQSKGPTGPKE 86
Db 61 GSQTHQVSLSKOPTSQSKGPTGPKE 86

RESULT 3
US-08-450-098-1
Sequence 1, Application US/08450098
Patent No. 5674980
GENERAL INFORMATION:
APPLICANT: FRANKEL, Alan
APPLICANT: PABO, Carl
APPLICANT: BARSOUM, James G.
APPLICANT: FAWELL, Stephen E.
APPLICANT: PEPINSKY, R. B.
TITLE OF INVENTION: TAT-DERIVED TRANSPORT POLYPEPTIDES
NUMBER OF SEQUENCES: 69
CORRESPONDENCE ADDRESS:
ADDRESSEE: FISH & NEAVE
STREET: 1251 Avenue of the Americas
CITY: New York
STATE: New York
COUNTRY: USA
ZIP: 10020
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patentin Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/450.098
FILING DATE: 25-MAY-1995
CLASSIFICATION: 514
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/235.403
FILING DATE: 28-APR-1994
APPLICATION NUMBER: US 07/934.375
FILING DATE: 21-AUG-1992
APPLICATION NUMBER: PCT/US93/07833
FILING DATE: 19-AUG-1993
APPLICATION NUMBER: US 07/454.450
FILING DATE: 21-DEC-1989
APPLICATION NUMBER: US 07/636.662
FILING DATE: 02-JAN-1991
APPLICATION NUMBER: US 08/158.015
FILING DATE: 24-NOV-1993
ATTORNEY/AGENT INFORMATION:
```

NAME: Haley Jr., James F.  
REGISTRATION NUMBER: 27,794  
REFERENCE/DOCKET NUMBER: B170 CIP 2  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: (212) 596-9000  
TELEFAX: (212) 596-9090  
TELEX: 14-8367  
INFORMATION FOR SEQ ID NO: 1:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 86 amino acids  
TYPE: amino acid  
TOPOLOGY: linear  
MOLECULE TYPE: protein  
ORIGINAL SOURCE:  
ORGANISM: human immunodeficiency virus  
STRAIN: type 1  
US-08-450-098-1

Query Match 86.6%; Score 478; DB 1; Length 86;  
Best Local Similarity 96.5%; Pred. No. 5.4e-46;  
Matches 83; Conservative 2; Mismatches 1; Indels 0; Gaps 0;

Qy 1 MEPVDPRLEPKHKGSPKQKTCNCKYKCCFHCQVCFITAAALGISYGRKKRRRRPPQ 60  
Db 1 MEPVDPRLEPKHKGSPKQKTCNCKYKCCFHCQVCFITAAALGISYGRKKRRRRPPQ 60

Qy 61 GSQTHVSLSKOPTSQSGEPTGPKE 86  
Db 61 GSQTHVSLSKOPTSQSGEPTGPKE 86

RESULT 4  
US-08-451-233-1  
; Sequence 1, Application US/08451233  
; Patent No. 5747641  
; GENERAL INFORMATION:  
; APPLICANT: FRANKEL, Alan  
; APPLICANT: PABO, Carl  
; APPLICANT: BARSOUM, James G.  
; APPLICANT: FAWELL, Stephen E.  
; APPLICANT: PEPINSKY, R. B.  
; TITLE OF INVENTION: TAT-DERIVED TRANSPORT POLYPEPTIDES  
; NUMBER OF SEQUENCES: 69  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: FISH & NEAVE  
; STREET: 1251 Avenue of the Americas  
; CITY: New York  
; STATE: New York  
; COUNTRY: USA  
; ZIP: 10020  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Floppy disk  
; COMPUTER: IBM PC compatible  
; OPERATING SYSTEM: PC-DOS/MS-DOS  
; SOFTWARE: PatentIn Release #1.0, Version #1.25  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/08/451,233  
; FILING DATE: 25-MAY-1995  
; CLASSIFICATION: 514  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: US 08/235,403  
; FILING DATE: 28-APR-1994  
; APPLICATION NUMBER: US 07/934,375  
; FILING DATE: 21-AUG-1992  
; APPLICATION NUMBER: US 07/098,766  
; FILING DATE: 28-JUL-1993  
; APPLICATION NUMBER: PCT/US93/07833  
; FILING DATE: 19-AUG-1993  
; APPLICATION NUMBER: US 07/454,450  
; FILING DATE: 21-DEC-1989  
; APPLICATION NUMBER: US 07/636,662  
; FILING DATE: 02-JAN-1991

APPLICATION NUMBER: US 08/158,015  
FILING DATE: 24-NOV-1993  
ATTORNEY/AGENT INFORMATION:  
NAME: Haley Jr., James F.  
REGISTRATION NUMBER: 27,794  
REFERENCE/DOCKET NUMBER: B170 CIP 2  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: (212) 596-9000  
TELEFAX: (212) 596-9090  
TELEX: 14-8367  
INFORMATION FOR SEQ ID NO: 1:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 86 amino acids  
TYPE: amino acid  
TOPOLOGY: linear  
MOLECULE TYPE: protein  
ORIGINAL SOURCE:  
ORGANISM: human immunodeficiency virus  
STRAIN: type 1  
US-08-451-233-1

Query Match 86.6%; Score 478; DB 1; Length 86;  
Best Local Similarity 96.5%; Pred. No. 5.4e-46;  
Matches 83; Conservative 2; Mismatches 1; Indels 0; Gaps 0;

Qy 1 MEPVDPRLEPKHKGSPKQKTCNCKYKCCFHCQVCFITAAALGISYGRKKRRRRPPQ 60  
Db 1 MEPVDPRLEPKHKGSPKQKTCNCKYKCCFHCQVCFITAAALGISYGRKKRRRRPPQ 60

Qy 61 GSQTHVSLSKOPTSQSGEPTGPKE 86  
Db 61 GSQTHVSLSKOPTSQSGEPTGPKE 86

RESULT 5  
US-08-450-236-1  
; Sequence 1, Application US/08450236  
; Patent No. 5804604  
; GENERAL INFORMATION:  
; APPLICANT: FRANKEL, Alan  
; APPLICANT: PABO, Carl  
; APPLICANT: BARSOUM, James G.  
; APPLICANT: FAWELL, Stephen E.  
; APPLICANT: PEPINSKY, R. B.  
; TITLE OF INVENTION: TAT-DERIVED TRANSPORT POLYPEPTIDES  
; NUMBER OF SEQUENCES: 69  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: FISH & NEAVE  
; STREET: 1251 Avenue of the Americas  
; CITY: New York  
; STATE: New York  
; COUNTRY: USA  
; ZIP: 10020  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Floppy disk  
; COMPUTER: IBM PC compatible  
; OPERATING SYSTEM: PC-DOS/MS-DOS  
; SOFTWARE: PatentIn Release #1.0, Version #1.25  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/08/450,236  
; FILING DATE: 25-MAY-1995  
; CLASSIFICATION: 514  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: US 08/235,403  
; FILING DATE: 28-APR-1994  
; APPLICATION NUMBER: US 07/934,375  
; FILING DATE: 21-AUG-1992  
; APPLICATION NUMBER: US 07/098,766  
; FILING DATE: 28-JUL-1993  
; APPLICATION NUMBER: PCT/US93/07833  
; FILING DATE: 19-AUG-1993  
; APPLICATION NUMBER: US 07/454,450

```

:      FILING DATE: 21-DEC-1989
:      APPLICATION NUMBER: US 07/636,862
:      FILING DATE: 02-JAN-1991
:      APPLICATION NUMBER: US 08/158,015
:      FILING DATE: 24-NOV-1993
:      ATTORNEY/AGENT INFORMATION:
:      NAME: Haley Jr., James F.
:      REGISTRATION NUMBER: 27,794
:      REFERENCE/DOCKET NUMBER: B170 CIP
:      TELECOMMUNICATION INFORMATION:
:      TELEPHONE: (212) 596-9000
:      TELEFAX: (212) 596-9090
:      TELEX: 14-8367
:      INFORMATION FOR SEQ ID NO: 1:
:      SEQUENCE CHARACTERISTICS:
:      LENGTH: 86 amino acids
:      TYPE: amino acid
:      TOPOLOGY: linear
:      MOLECULE TYPE: protein
:      ORIGINAL SOURCE:
:      ORGANISM: human immunodeficiency
:      STRAIN: type 1
:      US-08-450-236-1

```

Query Match 86.6%; Score 478; DB 1; Length 86;  
Best Local Similarity 96.5%; Pred. No. 5.4e-46;  
Matches 83; Conservative 2; Mismatches 1; Indels

QY	1	MEPVDPRLEPWHKPGSQSPKTTACTNCYCKKCCFHQVCVFITAA	60
		LGISYGRKKRRQRRPPQ	60
DB	1	MEPVDPRLEPWHKPGSQSPKTTACTNCYCKKCCFHQVCVFITK	60
		ALGISYGRKKRRQRRPPQ	60
QY	61	GSQTHQVLSLKQPTSSQSGKEPTGPKE	86
DB	61	GSQTHQVLSLKQPTSSQSGKEPTGPKE	86

```

RESULT      6
US-07-808-452-1
: Sequence 1, Application US/07808452
: Patent No. 6063612
: GENERAL INFORMATION:
: APPLICANT: Jayasena, Sumedha D.
: APPLICANT: Johnston, Brian H.
: TITLE OF INVENTION: Antiviral Reagents Based on
: TITLE OF INVENTION: RNA-Binding Proteins
: NUMBER OF SEQUENCES: 24
: CORRESPONDENCE ADDRESS:
: ADDRESSEE: Law Offices of Peter Dehlinger
: STREET: 350 Cambridge Avenue, Suite 300
: CITY: Palo Alto
: STATE: CA
: COUNTRY: USA
: ZIP: 94306
: COMPUTER READABLE FORM:
: MEDIUM TYPE: Floppy disk
: COMPUTER: IBM PC Compatible
: OPERATING SYSTEM: PC-DOS/MS-DOS
: SOFTWARE: Patent in Release #1.0, Version #1.25
: CURRENT APPLICATION DATA:
: APPLICATION NUMBER: US/07/808,452
: FILING DATE: 19911213
: CLASSIFICATION: 514
: ATTORNEY/AGENT INFORMATION:
: NAME: Fadian, Gary R.
: REGISTRATION NUMBER: 33,875
: REFERENCE/DOCKET NUMBER: 8255-008
: TELECOMMUNICATION INFORMATION:
: TELEPHONE: (415) 324-0880
: TELEFAX: (415) 324-0960
: INFORMATION FOR SEQ ID NO: 1:
: SEQUENCE CHARACTERISTICS:

```

```

; LENGTH: 86 amino acids
; TYPE: AMINO ACID
; TOPOLOGY: linear
; MOLECULE TYPE: protein
; HYPOTHETICAL: NO
; ANTI-SENSE: NO
; ORIGINAL SOURCE:
; INDIVIDUAL ISOLATE: the
; INDIVIDUAL ISOLATE: HIV
US-07-808-452-1

```

Query Match	86.6%	Score 478;	DB 3;	Length 86;
Best Local Similarity	96.5%	Pred. No. 5.4e-46;		
Matches	83;	Conservative	2;	Mismatches 1; Indels
Qy 1	MEPVDPRLEPMKHPGSOPKTACTNCYCKKCCFCQVCFTTAAALGYSYGRKKRRR			
Db 1	MEPVDPRLEPMKHPGSOPKTACTNCYCKKCCFCQVCFTTAAALGYSYGRKKRRR			
Qy 61	GSOTHVYSLSKQPTSSQSKSGEPTGPK			
Db 61	GSOTHVYSLSKQPTSSQSKSGEPTGPK			

RESULT 7  
 US-09-030-613-19  
 ; Sequence 19, Application US/09030613  
 ; Patent No. 6083706  
 ; GENERAL INFORMATION:  
 ; APPLICANT: Florlewicz, Robert Z.  
 ; APPLICANT: Baird, J. Andrew  
 ; TITLE OF INVENTION: INHIBITORS OF LEADERLESS PROTEIN EXPORT  
 ; NUMBER OF SEQUENCES: 36  
 ; CORRESPONDENCE ADDRESS:  
 ; ADDRESSEE: SEED and BERRY LLP  
 ; STREET: 6300 Columbia Center, 701 Fifth Avenue  
 ; CITY: Seattle  
 ; STATE: Washington  
 ; COUNTRY: USA  
 ; ZIP: 98104  
 ; COMPUTER READABLE FORM:  
 ; MEDIUM TYPE: Floppy disk  
 ; COMPUTER: IBM PC compatible  
 ; OPERATING SYSTEM: PC-DOS/MS-DOS  
 ; SOFTWARE: PatentIn Release #1.0, Version #1.30  
 ; CURRENT APPLICATION DATA:  
 ; APPLICATION NUMBER: US/09/030,613  
 ; FILING DATE: 25-FEB-1998  
 ; CLASSIFICATION:  
 ; ATTORNEY/AGENT INFORMATION:  
 ; NAME: No. 6083706tenburg Ph.D., Carol  
 ; REGISTRATION NUMBER: 39,317  
 ; REFERENCE/DOCKET NUMBER: 760100.418C1  
 ; TELECOMMUNICATION INFORMATION:  
 ; TELEPHONE: (206) 622-4900  
 ; TELEFAX: (206) 682-6031  
 ; INFORMATION FOR SEQ ID NO: 19:  
 ; SEQUENCE CHARACTERISTICS:  
 ; LENGTH: 86 amino acids  
 ; TYPE: amino acid  
 ; STRANDEDNESS:  
 ; TOPOLOGY: linear  
 US-09-030-613-19

	Query Match	86.6%	Score 478;	DB 3;	Length 86;
	Best Local Similarity	96.5%;	Prod. No. 5.4e-46;		
	Matches 83;	Conservative	2;	Mismatches 1;	Indels 0; Gaps 0;
Qy	1	MEVDPRLEPNKHPGQSPKPTACTNCYKCKCFHCQVCEITTAALGISYGRKKRRQRRRPQ	60		
Dp	1	MEVDPRLEPNKHPGQSPKPTACTNCYKCKCFHCQVCEITTAALGISYGRKKRRQRRRPQ	60		



Qy 61 GSQTHQVSLSKOPTSQSKGPTGPKE 86  
|||||  
Db 61 GSQTHQVSLSKOPTSQSKGPTGPKE 86

## RESULT 8

US-09-124-900-6  
; Sequence 6, Application US/09124900  
; Patent No. 6268484  
; GENERAL INFORMATION:  
; APPLICANT: KATINGER, Hermann  
; APPLICANT: BUCHACHER, Andrea  
; APPLICANT: ERNST, Wolfgang  
; APPLICANT: BALLAUN, Claudia  
; APPLICANT: PURTSCHER, Martin  
; APPLICANT: TRKOLA, Alexandra  
; APPLICANT: PREDL, Renate  
; APPLICANT: SCHMATZ, Christine  
; APPLICANT: KLIMA, Annelies  
; APPLICANT: STEINDL, Franz  
; APPLICANT: MUSTER, Thomas  
; TITLE OF INVENTION: HIV-Vaccines  
; FILE REFERENCE: 1939-112P  
; CURRENT APPLICATION NUMBER: US/09/124,900  
; CURRENT FILING DATE: 1998-07-30  
; PRIOR APPLICATION NUMBER: PCT/EP95/01481  
; PRIOR FILING DATE: 1995-04-19  
; NUMBER OF SEQ ID NOS: 11  
; SOFTWARE: PatentIn version 3.0  
; SEQ ID NO 6  
; LENGTH: 86  
; TYPE: PRT  
; ORGANISM: Human immunodeficiency virus type 1  
US-09-124-900-6

Query Match 86.6%; Score 478; DB 4; Length 86;  
Best Local Similarity 96.5%; Pred. No. 5.4e-46;  
Matches 83; Conservative 2; Mismatches 1; Indels 0; Gaps 0;

Qy 1 MEPVDPRLEPKHGPSQPKTACTNCKYCKKCCFHCQVCFITAAIGISYGRKKRRRRPPQ 60  
|||||  
Db 1 MEPVDPRLEPKHGPSQPKTACTNCKYCKKCCFHCQVCFITAAIGISYGRKKRRRRPPQ 60

Qy 61 GSQTHQVSLSKOPTSQSKGPTGPKE 86  
|||||  
Db 61 GSQTHQVSLSKOPTSQSKGPTGPKE 86

## RESULT 9

US-09-451-905-19  
; Sequence 19, Application US/09451905  
; Patent No. 6306613  
; GENERAL INFORMATION:  
; APPLICANT: Robert Z. Florkiewicz  
; APPLICANT: Andrew Baird  
; APPLICANT: Dale E. Warnock  
; TITLE OF INVENTION: MODULATORS OF LEADERLESS PROTEIN EXPORT  
; FILE REFERENCE: 200124.402C4  
; CURRENT APPLICATION NUMBER: US/09/451,905  
; CURRENT FILING DATE: 1999-12-01  
; NUMBER OF SEQ ID NOS: 48  
; SOFTWARE: FastSeq for Windows Version 4.0  
; SEQ ID NO 19  
; LENGTH: 86  
; TYPE: PRT  
; ORGANISM: Homo sapien  
US-09-451-905-19

Query Match 86.6%; Score 478; DB 4; Length 86;

Best Local Similarity 96.5%; Pred. No. 5.4e-46;  
Matches 83; Conservative 2; Mismatches 1; Indels 0; Gaps 0;  
Qy 1 MEPVDPRLEPKHGPSQPKTACTNCKYCKKCCFHCQVCFITAAIGISYGRKKRRRRPPQ 60  
|||||  
Db 1 MEPVDPRLEPKHGPSQPKTACTNCKYCKKCCFHCQVCFITAAIGISYGRKKRRRRPPQ 60

Qy 61 GSQTHQVSLSKOPTSQSKGPTGPKE 86  
|||||  
Db 61 GSQTHQVSLSKOPTSQSKGPTGPKE 86

## RESULT 10

US-08-235-403-1  
; Sequence 1, Application US/08235403  
; Patent No. 6316003  
; GENERAL INFORMATION:  
; APPLICANT: FRANKEL, Alan  
; APPLICANT: PABO, Carl  
; APPLICANT: BARSOUM, James G.  
; APPLICANT: FAWELL, Stephen E.  
; APPLICANT: PEPINSKY, R. B.  
; TITLE OF INVENTION: TAT-DERIVED TRANSPORT POLYPEPTIDES  
; NUMBER OF SEQUENCES: 69  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: FISH & NEAVE  
; STREET: 1251 Avenue of the Americas  
; CITY: New York  
; STATE: New York  
; COUNTRY: USA  
; ZIP: 10020  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Floppy disk  
; COMPUTER: IBM PC compatible  
; OPERATING SYSTEM: PC-DOS/MS-DOS  
; SOFTWARE: PatentIn Release #1.0, Version #1.25  
; CURRENT APPLICATION NUMBER: US/08/235,403  
; FILING DATE:  
; CLASSIFICATION: 514  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: US 07/934,375  
; FILING DATE: 21-AUG-1992  
; APPLICATION NUMBER: US 07/098,766  
; FILING DATE: 28-JUL-1993  
; APPLICATION NUMBER: PCT/US93/07833  
; FILING DATE: 19-AUG-1993  
; APPLICATION NUMBER: US 07/454,450  
; FILING DATE: 21-DEC-1989  
; APPLICATION NUMBER: US 07/636,662  
; FILING DATE: 02-JAN-1991  
; APPLICATION NUMBER: US 08/158,015  
; FILING DATE: 24-NOV-1993  
; ATTORNEY/AGENT INFORMATION:  
; NAME: Haley Jr., James F.  
; REGISTRATION NUMBER: 27,794  
; REFERENCE/DOCKET NUMBER: B170 CIP 2  
; TELECOMMUNICATION INFORMATION:  
; TELEPHONE: (212) 596-9000  
; TELEFAX: (212) 596-9090  
; TELEX: 14-8367  
; INFORMATION FOR SEQ ID NO: 1:  
; SEQUENCE CHARACTERISTICS:  
; LENGTH: 86 amino acids  
; TYPE: amino acid  
; TOPOLOGY: linear  
; MOLECULE TYPE: protein  
; ORIGINAL SOURCE:  
; ORGANISM: human immunodeficiency virus  
; STRAIN: type 1  
US-08-235-403-1

```
Query Match      86.6%; Score 478; DB 4; Length 86;
Best Local Similarity 96.5%; Pred. No. 5.4e-46;
Matches 83; Conservative 2; Mismatches 1; Indels 0; Gaps 0;

Qy 1 MEPVDPRLPEPKHPSQPKTACTNCKKCCFHCQVCFITAAALGISYGRKKRRQRRRPQ 60
Db 1 MEPVDPRLPEPKHPSQPKTACTNCKKCCFHCQVCFITKALGISYGRKKRRQRRRPQ 60

Qy 61 GSQTHQVSLSKOPTSQSKGEPTGPKE 86
Db 61 GSQTHQVSLSKOPTSQSKGEPTGPKE 86

RESULT 11
PCT-US92-10770-1
; Sequence 1, Application PC/TUS9210770
; GENERAL INFORMATION:
; APPLICANT: Jayasena, Suredha D.
; APPLICANT: Johnston, Brian H.
; TITLE OF INVENTION: Antiviral Reagents Based on
; TITLE OF INVENTION: RNA-Binding Proteins
; NUMBER OF SEQUENCES: 24
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: SRI International
; STREET: 333 Ravenswood Avenue
; CITY: Menlo Park
; STATE: CA
; COUNTRY: USA
; ZIP: 94025
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: PCT/US92/10770
; FILING DATE: 19921211
; CLASSIFICATION:
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 07/808,452
; FILING DATE: 13-DEC-1991
; ATTORNEY/AGENT INFORMATION:
; NAME: Fabian, Gary R.
; REGISTRATION NUMBER: 33,875
; REFERENCE/DOCKET NUMBER: P-2962
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (415) 859-4550
; TELEFAX: (415) 859-3880
; INFORMATION FOR SEQ ID NO: 1:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 86 amino acids
; TYPE: AMINO ACID
; TOPOLOGY: linear
; MOLECULE TYPE: protein
; HYPOTHETICAL: NO
; ANTI-SENSE: NO
; ORIGINAL SOURCE:
; INDIVIDUAL ISOLATE: the sequence of the TAT protein of
; INDIVIDUAL ISOLATE: HIV-1
PCT-US92-10770-1

Query Match      86.6%; Score 478; DB 5; Length 86;
Best Local Similarity 96.5%; Pred. No. 5.4e-46;
Matches 83; Conservative 2; Mismatches 1; Indels 0; Gaps 0;

Qy 1 MEPVDPRLPEPKHPSQPKTACTNCKKCCFHCQVCFITAAALGISYGRKKRRQRRRPQ 60
Db 1 MEPVDPRLPEPKHPSQPKTACTNCKKCCFHCQVCFITKALGISYGRKKRRQRRRPQ 60

Qy 61 GSQTHQVSLSKOPTSQSKGEPTGPKE 86
Db 61 GSQTHQVSLSKOPTSQSKGEPTGPKE 86
```

```
RESULT 12
PCT-US95-06077-2
; Sequence 2, Application PC/TUS9506077
; GENERAL INFORMATION:
; APPLICANT: Immunobiology Research, Institute Inc.
; TITLE OF INVENTION: Vaccine Interdiction of Extracellular
; TITLE OF INVENTION: Transactivating Proteins of Human Immunodeficiency Virus
; TITLE OF INVENTION: and Other Chronically Infecting Viruses Employing Similar
; TITLE OF INVENTION: InterCellular Transactivating Strategies
; NUMBER OF SEQUENCES: 38
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Howson and Howson
; STREET: Spring House Corporate Cntr, PO Box 457
; CITY: Spring House
; STATE: Pennsylvania
; COUNTRY: USA
; ZIP: 19477
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: PCT/US95/06077
; FILING DATE:
; CLASSIFICATION:
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/247,991
; FILING DATE: 23-MAY-1994
; ATTORNEY/AGENT INFORMATION:
; NAME: Bak, Mary E.
; REGISTRATION NUMBER: 31,215
; REFERENCE/DOCKET NUMBER: IRI44PCT
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 215-540-9200
; TELEFAX: 215-540-5818
; INFORMATION FOR SEQ ID NO: 2:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 86 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: protein
PCT-US95-06077-2

Query Match      86.6%; Score 478; DB 5; Length 86;
Best Local Similarity 96.5%; Pred. No. 5.4e-46;
Matches 83; Conservative 2; Mismatches 1; Indels 0; Gaps 0;

Qy 1 MEPVDPRLPEPKHPSQPKTACTNCKKCCFHCQVCFITAAALGISYGRKKRRQRRRPQ 60
Db 1 MEPVDPRLPEPKHPSQPKTACTNCKKCCFHCQVCFITKALGISYGRKKRRQRRRPQ 60

Qy 61 GSQTHQVSLSKOPTSQSKGEPTGPKE 86
Db 61 GSQTHQVSLSKOPTSQSKGEPTGPKE 86

RESULT 13
US-08-505-210-1
; Sequence 1, Application US/08505210
; Patent No. 5981258
; GENERAL INFORMATION:
; APPLICANT: MEHTALI, Majid
; APPLICANT: GUSS, Tania
; TITLE OF INVENTION: COMPOSITION OF TRANS-DOMINANT VARIANTS
; TITLE OF INVENTION: OF VIRAL PROTEINS FOR OBTAINING AN ANTIVIRAL EFFECT
; NUMBER OF SEQUENCES: 7
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: BURNS, DOANE, SWECKER & MATHIS, L.L.P.
; STREET: 1737 King Street, Suite 500
```





RESULT 19  
US-09-113-921-1  
; Sequence 1, Application US/09113921  
; Patent No. 6193981  
; GENERAL INFORMATION:  
; APPLICANT: Goldstein, Gideon  
; TITLE OF INVENTION: Methods and Compositions for Impairing  
; TITLE OF INVENTION: Multiplication of HIV-1  
; NUMBER OF SEQUENCES: 124  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: Howson and Howson  
; STREET: Spring House Corporate Cntr., P.O. Box 457  
; CITY: Spring House  
; STATE: PA  
; COUNTRY: USA  
; ZIP: 19477  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Floppy disk  
; COMPUTER: IBM PC compatible  
; OPERATING SYSTEM: PC-DOS/MS-DOS  
; SOFTWARE: PatentIn Release #1.0, Version #1.30  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/09/113.921  
; FILING DATE:  
; CLASSIFICATION:  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: US 08/893,853  
; FILING DATE: 11-JUL-1997  
; ATTORNEY/AGENT INFORMATION:  
; NAME: Bak, Mary E.  
; REGISTRATION NUMBER: 31,215  
; REFERENCE/DOCKET NUMBER: GGP2AUSA  
; TELECOMMUNICATION INFORMATION:  
; TELEPHONE: 215-540-9200  
; TELEFAX: 215-540-5818  
; INFORMATION FOR SEQ ID NO: 1:  
; SEQUENCE CHARACTERISTICS:  
; LENGTH: 72 amino acids  
; TYPE: amino acid  
; STRANDEDNESS:  
; TOPOLOGY: linear  
; MOLECULE TYPE: protein  
US-09-113-921-1  
  
Query Match 69.6%; Score 384; DB 4; Length 72;  
Best Local Similarity 93.1%; Pred. No. 1.le-35;  
Matches 67; Conservative 0; Mismatches 5; Indels 0; Gaps 0;  
  
QY 1 MEPVDPRLPKHPSQPKTACTNCKKCCFHCQVCFITAAIGISYGRKKRRRPPQ 60  
Db 1 MEPVDPRLPKHPSQPKTACTNCKKCCFHCQVCFITAAIGISYGRKKRRRPPQ 60  
  
QY 61 GSQTHQVSLSKQ 72  
Db 61 DSQTHQVSLSKQ 72  
  
RESULT 20  
US-09-128A-27  
; Sequence 27, Application US/08094128A  
; Patent No. 5595884  
; GENERAL INFORMATION:  
; APPLICANT: BARSOUM, James G.  
; APPLICANT: ANDROPHY, Elliot J.  
; TITLE OF INVENTION: PAPILLOMAVIRUS E2 TRANS-ACTIVATION  
; TITLE OF INVENTION: REPRESSORS  
; NUMBER OF SEQUENCES: 33  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: FISH & NEAVE  
; STREET: 1251 Avenue of the Americas  
; CITY: New York

STATE: New York  
COUNTRY: USA  
ZIP: 10020  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: PatentIn Release #1.0, Version #1.25  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/094.128A  
FILING DATE: 24-SEP-1993  
CLASSIFICATION: 517  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: US 07/646,998  
FILING DATE: 28-JAN-1991  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: PCT/US92/00652  
FILING DATE: 28-JAN-1992  
ATTORNEY/AGENT INFORMATION:  
NAME: Haley Jr., James F.  
REGISTRATION NUMBER: 27,794  
REFERENCE/DOCKET NUMBER: B156CIP  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: (212) 596-9000  
TELEFAX: (212) 596-9090  
INFORMATION FOR SEQ ID NO: 27:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 312 amino acids  
TYPE: amino acid  
TOPOLOGY: linear  
MOLECULE TYPE: protein  
US-08-094-128A-27  
  
Query Match 65.3%; Score 360.5; DB 1; Length 312;  
Best Local Similarity 70.2%; Pred. No. 2.3e-32;  
Matches 66; Conservative 6; Mismatches 13; Indels 9; Gaps 2;  
  
QY 1 MEPVDPRLPKHPSQPKTACTNCKKCCFHCQVCFITAAIGISYGRKKRRRPPQ 60  
Db 1 MEPVDPRLPKHPSQPKTACTNCKKCCFHCQVCFITAAIGISYGRKKRRRPPQ 60  
  
QY 61 GSQTHQVSLSKQTSQSGKGTGPKETSGHH 91  
Db 61 GSSMAGAGRIYYSR-----FGDEAARFTTGHY 88  
  
RESULT 21  
US-08-455-674-27  
; Sequence 27, Application US/08455674  
; Patent No. 5616559  
; GENERAL INFORMATION:  
; APPLICANT: BARSOUM, James G.  
; APPLICANT: ANDROPHY, Elliot J.  
; TITLE OF INVENTION: PAPILLOMAVIRUS E2 TRANS-ACTIVATION  
; TITLE OF INVENTION: REPRESSORS  
; NUMBER OF SEQUENCES: 33  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: FISH & NEAVE  
; STREET: 1251 Avenue of the Americas  
; CITY: New York  
; STATE: New York  
; COUNTRY: USA  
; ZIP: 10020  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Floppy disk  
; COMPUTER: IBM PC compatible  
; OPERATING SYSTEM: PC-DOS/MS-DOS  
; SOFTWARE: PatentIn Release #1.0, Version #1.25  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/08/455,674  
; FILING DATE: 31-MAY-1995  
; CLASSIFICATION: 514

```

; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/094,128
; FILING DATE: 24-SEP-1993
; APPLICATION NUMBER: US 07/646,998
; FILING DATE: 28-JAN-1991
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: PCT/US92/006552
; FILING DATE: 28-JAN-1992
; ATTORNEY/AGENT INFORMATION:
; NAME: Haley Jr., James F.
; REGISTRATION NUMBER: 27,794
; REFERENCE/DOCKET NUMBER: B156CIP
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (212) 596-9000
; TELEFAX: (212) 596-9030
; INFORMATION FOR SEQ ID NO: 27:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 312 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: protein
; US-08-455-674-27

```

```

Query Match      55.3%; Score 360.5; DB 1; Length 312;
Best Local Similarity 70.2%; Pred. No. 2.3e-32;
Matches 66; Conservative 6; Mismatches 13; Indels 9; Gaps 2;

QY 1 MEVDPRLPEWKHPGSPKACTNICYKCKCFHCQVCFITAAALGISYGRKKRRORRPPQ 60
      |||||
Db 1 MEVDPRLPEWKHPGSPKACTNICYKCKCFHCQVCFITKALGISYGRKKRRORRPPQ 60
      |||||

QY 61 GSQ---THQVSLSKQPTSGKQPTGPKETSGHH 91
      ||| :||:
Db 61 GSSMAGAGRTIYSR-----FGDEAARFSTTGHV 88

```

RESULT 22  
US-08-455-992-27  
; Sequence 27, Application US/08455992  
; Patent No. 5656599  
; GENERAL INFORMATION:  
; APPLICANT: BARSOUM, James G.  
; APPLICANT: ANDROPHY, Elliot J.  
; TITLE OF INVENTION: PAPILLOMAVIRUS E2 TRANS-ACTIVATION  
; TITLE OF INVENTION: REPRESSORS  
; NUMBER OF SEQUENCES: 33  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: FISH & NEAVE  
; STREET: 1251 Avenue of the Americas  
; CITY: New York  
; STATE: New York  
; COUNTRY: USA  
; ZIP: 10020  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Floppy disk  
; COMPUTER: IBM PC Compatible  
; OPERATING SYSTEM: PC-DOS/MS-DOS  
; SOFTWARE: Patent In Release #1.0, Version #1.25  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/08/455.992  
; FILING DATE: 31-MAY-1995  
; CLASSIFICATION: 530  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: US 08/094,128  
; FILING DATE: 24-SEP-1993  
; APPLICATION NUMBER: US 07/646,998  
; FILING DATE: 28-JAN-1991  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: PCT/US92/00652  
; FILING DATE: 28-JAN-1992  
; ATTORNEY/AGENT INFORMATION:  
; NAME: Haley Jr., James F.

```

; REGISTRATION NUMBER: 27,794
; REFERENCE/DOCKET NUMBER: B156CIP
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (212) 596-9000
; TELEFAX: (212) 596-9090
; INFORMATION FOR SEQ ID NO: 27:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 312 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: protein
US-08-455-992-27

Query Match          65.3%; Score 360.5; DB 1; Length
Best Local Similarity 70.2%; Pred. No. 2.3e-32;
Matches 56; Conservative 6; Mismatches 13; Indels

QY      1 MEVDPRLPEWPKHGSQPQTACTNCYCKKCFHCQCVCFTTAALGISTYGRKKR
        |||||
Db       1 MEVDPRLPEWPKHGSQPQTACTNCYCKKCFHCQCVCFTTAKLGIYSTYGRKKR
        |||||

QY      61 GSQ---THQVSLSLKQPTSQSKGEPTGPRTSGHH 91
        || : :: | : : |||
Db       61 GSSMAGAGRIYYSR-----FGDEAARFTTGHY 88

RESULT 23
US-08-455-972-27
; Sequence 27, Application US/08455972
; Patent No. 5667965
; GENERAL INFORMATION:
; APPLICANT: BARSOOM, James G.
; APPLICANT: ANDROPHY, Elliot J.
; TITLE OF INVENTION: PAPILLOMAVIRUS E2 TRANS-ACTIVATION
; TITLE OF INVENTION: REPRESSORS
; NUMBER OF SEQUENCES: 33
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: FISH & NEAVE
; STREET: 1251 Avenue of the Americas
; CITY: New York
; STATE: New York
; COUNTRY: USA
; ZIP: 10020
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/RMS-DOS
; SOFTWARE: Patent In Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/455,972
; FILING DATE: 31-MAY-1995
; CLASSIFICATION: 435
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/094,128
; FILING DATE: 24-SEP-1993
; APPLICATION NUMBER: US 07/646,998
; FILING DATE: 28-JAN-1991
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: PCT/US92/00652
; FILING DATE: 28-JAN-1992
; ATTORNEY/AGENT INFORMATION:
; NAME: Haley Jr., James F.
; REGISTRATION NUMBER: 27,794
; REFERENCE/DOCKET NUMBER: B156CIP
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (212) 596-9000
; TELEFAX: (212) 596-9090
; INFORMATION FOR SEQ ID NO: 27:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 312 amino acids
; TYPE: amino acid
; TOPOLOGY: linear

```

MOLECULE TYPE: protein

US-08-455-972-27

Query Match 65.3%; Score 360.5; DB 1; Length 312;  
Best Local Similarity 70.2%; Pred. No. 2.3e-32;  
Matches 66; Conservative 6; Mismatches 13; Indels 9; Gaps 2;

Qy 1 MEPVDPRLPEPKHPSQPKTACTNCKYKCCFHCQVCFITTAALGISYGRKKRRRPPQ 60  
Db 1 MEPVDPRLPEPKHPSQPKTACTNCKYKCCFHCQVCFITTAALGISYGRKKRRRPPQ 60  
Qy 61 GSQ---THQVLSKQPTSGKGTGPKETSQHH 91  
Db 61 GSSMAGAGRIIYSR-----FGDEAARFSTTGHY 88

RESULT 24

PCT-US92-00652-27  
Sequence 27, Application PC/TUS9200652  
GENERAL INFORMATION:  
APPLICANT: BIOGEN, INC.  
APPLICANT: NEW ENGLAND MEDICAL CENTER,  
APPLICANT: HOSPITALS, INC.  
APPLICANT: BARSOUM, James G. (US only)  
APPLICANT: ANDROPHY, Elliot J. (US only)  
TITLE OF INVENTION: PAPILLOMAVIRUS E2 TRANS-ACTIVATION  
TITLE OF INVENTION: REPRESSORS  
NUMBER OF SEQUENCES: 30  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: FISH & NEAVE  
STREET: 875 Third Avenue  
CITY: New York  
STATE: New York  
COUNTRY: USA  
ZIP: 10022-6250

COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: PatentIn Release #1.0, Version #1.25  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: PCT/US92/00652  
FILING DATE: 19920128  
CLASSIFICATION: 435  
ATTORNEY/AGENT INFORMATION:  
NAME: Haley Jr., James F.  
REGISTRATION NUMBER: 27,794  
REFERENCE/DOCKET NUMBER: B156CIP  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: (212)715-0600  
TELEFAX: (212)715-0673  
TELEX: 14-8367

INFORMATION FOR SEQ ID NO: 27:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 312 amino acids  
TYPE: AMINO ACID  
TOPOLOGY: linear  
MOLECULE TYPE: protein  
PCT-US92-00652-27

Query Match 65.3%; Score 360.5; DB 5; Length 312;  
Best Local Similarity 70.2%; Pred. No. 2.3e-32;  
Matches 66; Conservative 6; Mismatches 13; Indels 9; Gaps 2;

Qy 1 MEPVDPRLPEPKHPSQPKTACTNCKYKCCFHCQVCFITTAALGISYGRKKRRRPPQ 60  
Db 1 MEPVDPRLPEPKHPSQPKTACTNCKYKCCFHCQVCFITTAALGISYGRKKRRRPPQ 60  
Qy 61 GSQ---THQVLSKQPTSGKGTGPKETSQHH 91  
Db 61 GSSMAGAGRIIYSR-----FGDEAARFSTTGHY 88

RESULT 25

US-07-910-867B-2  
Sequence 2, Application US/07910867B  
Patent No. 5597895  
GENERAL INFORMATION:  
APPLICANT: Gaynor, Richard B.  
APPLICANT: Garcia, Joseph A.  
APPLICANT: Harrich, David  
TITLE OF INVENTION: Transdominant Tat Mutants and Uses  
NUMBER OF SEQUENCES: 22  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: Arnold, White & Durkee  
STREET: P.O. Box 4433  
CITY: Houston  
STATE: Texas  
COUNTRY: US  
ZIP: 77210

COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: PatentIn Release #1.0, Version #1.30  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/07/910,867B  
FILING DATE: 02-JUL-1992  
CLASSIFICATION: 435  
ATTORNEY/AGENT INFORMATION:  
NAME: Mayfield, Denise L.  
REGISTRATION NUMBER: 33,732  
REFERENCE/DOCKET NUMBER: UTSD:263/MAY  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: 512/418-3000  
TELEFAX: 512/474-7577  
TELEX: N/A

INFORMATION FOR SEQ ID NO: 2:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 72 amino acids  
TYPE: amino acid  
STRANDEDNESS: single  
TOPOLOGY: linear  
MOLECULE TYPE: peptide  
US-07-910-867B-2

Query Match 63.6%; Score 351; DB 1; Length 72;  
Best Local Similarity 83.3%; Pred. No. 4.8e-32;  
Matches 60; Conservative 1; Mismatches 11; Indels 0; Gaps 0;

Qy 1 MEPVDPRLPEPKHPSQPKTACTNCKYKCCFHCQVCFITTAALGISYGRKKRRRPPQ 60  
Db 1 MEPVDPRLPEPKHPSQPKTACTNCKYKCCFHCQVCFITTAALGISYGRKKRRRPPQ 60

Qy 61 GSQTHQVLSKQ 72  
Db 61 DSQTHQVLSKQ 72

RESULT 26

US-07-910-867B-5  
Sequence 5, Application US/07910867B  
Patent No. 5597895  
GENERAL INFORMATION:  
APPLICANT: Gaynor, Richard B.  
APPLICANT: Garcia, Joseph A.  
APPLICANT: Harrich, David  
TITLE OF INVENTION: Transdominant Tat Mutants and Uses  
NUMBER OF SEQUENCES: 22  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: Arnold, White & Durkee

STREET: P.O. Box 4433  
CITY: Houston  
STATE: Texas  
COUNTRY: US  
ZIP: 77210  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: PatentIn Release #1.0, Version #1.30  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/07/910,867B  
FILING DATE: 02-JUL-1992  
CLASSIFICATION: 435  
ATTORNEY/AGENT INFORMATION:  
NAME: Mayfield, Denise L.  
REGISTRATION NUMBER: 33,732  
REFERENCE/DOCKET NUMBER: UTSD:263/MAY  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: 512/418-3000  
TELEFAX: 512/474-7577  
TELEX: N/A  
INFORMATION FOR SEQ ID NO: 5:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 72 amino acids  
TYPE: amino acid  
STRANDEDNESS: single  
TOPOLOGY: linear  
MOLECULE TYPE: peptide  
US-07-910-867B-5

Query Match 59.88; Score 330; DB 1; Length 72;  
Best Local Similarity 79.28; Pred. No. 1e-29;  
Matches 57; Conservative 1; Mismatches 14; Indels 0; Gaps 0;

QY 1 MEPVDPRLPEPKHGPSQPKTACTNCKYKCCFHCQVCFITAAALGISYGRKKRRRRPPQ 60  
||||| ||||||| ||||||| ||||||| ||||||| ||||||| ||||||| ||  
Db 1 MEPVDPNLEPKHGPSQRTACNNCKYKCCFHCYACFTRKGLGISYGRKKRAGGAPQ 60  
||||| ||||||| ||||||| ||||||| ||||||| ||||||| ||||||| ||  
QY 61 GSQTHQVLSKQ 72  
||||| ||||||  
Db 61 DSQTHQASLSKQ 72

RESULT 27  
US-07-910-867B-3  
Sequence 3, Application US/07910867B  
Patent No. 5597895  
GENERAL INFORMATION:  
APPLICANT: Gaynor, Richard B.  
APPLICANT: Garcia, Joseph A.  
APPLICANT: Harrich, David  
TITLE OF INVENTION: Transdominant Tat Mutants and Uses  
TITLE OF INVENTION: Theof  
NUMBER OF SEQUENCES: 22  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: Arnold, White & Durkee  
STREET: P.O. Box 4433  
CITY: Houston  
STATE: Texas  
COUNTRY: US  
ZIP: 77210  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: PatentIn Release #1.0, Version #1.30  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/07/910,867B  
FILING DATE: 02-JUL-1992  
CLASSIFICATION: 435  
ATTORNEY/AGENT INFORMATION:

NAME: Mayfield, Denise L.  
REGISTRATION NUMBER: 33,732  
REFERENCE/DOCKET NUMBER: UTSD:263/MAY  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: 512/418-3000  
TELEFAX: 512/474-7577  
TELEX: N/A  
INFORMATION FOR SEQ ID NO: 3:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 72 amino acids  
TYPE: amino acid  
STRANDEDNESS: single  
TOPOLOGY: linear  
MOLECULE TYPE: peptide  
US-07-910-867B-3

Query Match 58.78; Score 324; DB 1; Length 72;  
Best Local Similarity 77.88; Pred. No. 4.7e-29;  
Matches 56; Conservative 1; Mismatches 15; Indels 0; Gaps 0;

QY 1 MEPVDPRLPEPKHGPSQPKTACTNCKYKCCFHCQVCFITAAALGISYGRKKRRRRPPQ 60  
||||| ||||||| ||||||| ||||||| ||||||| ||||||| ||||||| ||  
Db 1 MEPVDPNLEPKHGPSQRTACNNCKYKCCFHCYACFTRKGLGISYGRKKRAGGAPQ 60  
||||| ||||||| ||||||| ||||||| ||||||| ||||||| ||||||| ||  
QY 61 GSQTHQVLSKQ 72  
||||| ||||||  
Db 61 DSQTHQASLSKQ 72

RESULT 28  
US-07-910-867B-1  
Sequence 1, Application US/07910867B  
Patent No. 5597895  
GENERAL INFORMATION:  
APPLICANT: Gaynor, Richard B.  
APPLICANT: Garcia, Joseph A.  
APPLICANT: Harrich, David  
TITLE OF INVENTION: Transdominant Tat Mutants and Uses  
TITLE OF INVENTION: Theof  
NUMBER OF SEQUENCES: 22  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: Arnold, White & Durkee  
STREET: P.O. Box 4433  
CITY: Houston  
STATE: Texas  
COUNTRY: US  
ZIP: 77210  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: PatentIn Release #1.0, Version #1.30  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/07/910,867B  
FILING DATE: 02-JUL-1992  
CLASSIFICATION: 435  
ATTORNEY/AGENT INFORMATION:  
NAME: Mayfield, Denise L.  
REGISTRATION NUMBER: 33,732  
REFERENCE/DOCKET NUMBER: UTSD:263/MAY  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: 512/418-3000  
TELEFAX: 512/474-7577  
TELEX: N/A  
INFORMATION FOR SEQ ID NO: 1:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 72 amino acids  
TYPE: amino acid  
STRANDEDNESS: single  
TOPOLOGY: linear  
MOLECULE TYPE: peptide  
US-07-910-867B-1





```

Query Match      50.0%; Score 276; DB 1; Length 56;
Best Local Similarity 76.4%; Pred. No. 7.1e-24;
Matches 55; Conservative 0; Mismatches 1; Indels 16; Gaps 1;

QY 1 MEVDPRLEPWKHPGSOPTACTNCKCKCFHCQVGFITAAAGISYGRKKRRRRPPQ 60
    |||||
Db 1 MEVDPRLEPWKHPGSOPTACTNCKCKCFHCQVGFITAAAGISYGRKKRRRRPPQ 60
    |||||
QY 61 GSQTHQVSLSKQ 72
    |||||
Db 45 GSQTHQVSLSKQ 56
    |||||

```

Search completed: August 26, 2002, 08:10:41  
Job time: 130 sec

GenCore version 4.5  
Copyright (c) 1993 - 2000 Compugen Ltd.

OM protein - protein search, using sw model

Run on: August 26, 2002, 08:13:23 ; Search time 50.45 seconds  
(without alignments)  
180.941 Million cell updates/sec

Title: US-09-509-239-23  
Perfect score: 552  
Sequence: 1 MEPVDPRLPWHKPGSQPKT.....QSKGPTGPKETSGHHHHH 95

Scoring table: BLOSUM62  
Gapop 10.0 , Gapext 0.5

Searched: 283138 seqs, 96089334 residues

Total number of hits satisfying chosen parameters: 283138

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

1: pir1.\*

2: pir2.\*

3: pir3.\*

4: pir4.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

#### SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	478	86.6	86	2 S33982	trans-activating t
2	470	85.1	95	1 TNLJ12	trans-activating t
3	462	83.7	86	2 A25700	trans-activating t
4	429	77.7	101	2 T09446	tat protein - huma
5	426	77.2	101	1 E44001	trans-activating t
6	403	73.0	86	2 JC5591	transactivator pro
7	400	72.5	86	1 TNLJZR	trans-activating t
8	400	72.5	86	2 S34381	tat protein - huma
9	395	71.6	86	1 TNLJND	trans-activating t
10	387.5	70.2	87	2 T01665	tat protein - huma
11	364	65.9	72	1 TNLJH4	trans-activating t
12	353	63.9	71	2 T09384	trans-activating t
13	315.5	57.2	100	1 TNLJSI	trans-activating t
14	175.5	31.8	130	1 TNLJGG	trans-activating t
15	170.5	30.9	130	1 TNLJST	trans-activating t
16	170.5	30.9	130	2 S53096	tat protein - huma
17	163.5	29.6	130	1 TNLJG2	trans-activating t
18	161.5	29.3	133	1 TNLJCA	trans-activating t
19	161	29.2	130	2 S12157	trans-activating t
20	151	27.4	129	1 TNLJG3	trans-activating t
21	149	27.0	106	1 TNLJS2	trans-activating t
22	148.5	26.9	96	2 S08440	trans-activating t
23	142	25.7	100	1 TNLJG4	trans-activating t
24	139.5	25.3	119	2 A46356	tat protein - simi
25	138.5	25.1	132	2 T11564	tat protein - simi
26	133.5	24.2	131	2 T11557	trans-activating t
27	123	22.3	116	1 A48344	trans-activating t
28	114.5	20.7	145	1 TNLJBT	trans-activating t
29	112	20.3	73	2 S46349	trans-activating t

30 88 15.9 1001 2 T28897  
31 85.5 15.5 617 2 T23197  
32 85 15.4 1477 2 T13797  
33 82.5 14.9 346 2 F96785  
34 82.5 14.9 3942 2 T42730  
35 78 14.1 439 2 T46375  
36 77 13.9 174 2 G84600  
37 77 13.9 944 1 S48821  
38 76 13.8 354 2 S39406  
39 76 13.8 355 2 S35345  
40 76 13.8 1802 2 T00020  
41 76 13.8 2273 2 T46477  
42 76 13.8 2424 2 T46480  
43 75.5 13.7 3938 2 T42761  
44 75 13.6 355 2 T56547  
45 75 13.6 1069 2 T43280

#### ALIGNMENTS

RESULT 1

S33982

trans-activating transcription regulator - human immunodeficiency virus type 1, HIV-1  
C:Species: human immunodeficiency virus type 1, HIV-1

C:Date: 06-Oct-1994 #sequence\_revision 23-Feb-1996 #text\_change 20-Sep-1999

C:Accession: S33982; S26385; S19864

R:Carlini, F.

submitted to the EMBL Data Library, November 1991

A:Reference number: S33979

A:Accession: S33982

A:Molecule type: DNA

A:Residues: 1-86 <CAR>

A:Cross-references: EMBL:211530; NID:960192; PIDN:CAA7625.1; PID:960196

R:Stiderovski, D.P.; Matsuyama, T.; Frigerio, E.; Chui, S.; Min, X.; Erfle, H.; Sumner

Nucleic Acids Res. 20, 5311-5320, 1992

A:Title: Random mutagenesis of the human immunodeficiency virus type-1 trans-activator

A:Reference number: S26385; MUID:93065196

A:Accession: S26385

A:Molecule type: nucleic acid

A:Residues: 1-86 <SID>

A:Cross-references: EMBL:X64650; NID:960144; PIDN:CAA45921.1; PID:960145

C:Genetics:

A:Gene: tat

A:Introns: 72/2

C:Superfamily: AIDS trans-activating transcription regulator

C:Keywords: AIDS; immunodeficiency

Query Match 86.6%; Score 478; DB 2; Length 86;

Best Local Similarity 96.5%; Pred. No. 2e-37; Mismatches 2; Indels 0; Gaps 0;

Matches 83; Conservative 2; Mismatches 1; Indels 0; Gaps 0;

QY 1 MEPVDPRLPWHKPGSQPKTACTNCYCKKCCFHCQVCFTALGISYGRKKRRRRPPQ 60

|||||

Db 1 MEPVDPRLPWHKPGSQPKTACTNCYCKKCCFHCQVCFTALGISYGRKKRRRRPPQ 60

QY 61 GSQTHQVSLSKOPTSOSKGEPTGPKE 86

|||||

Db 61 GSQTHQVSLSKOPTSOSKGEPTGPKE 86

RESULT 2

TNLJ12

trans-activating transcription regulator - human immunodeficiency virus type 1 (isola

C:Species: human immunodeficiency virus type 1, HIV-1

A>Note: host Homo sapiens (man)

C:Date: 04-Dec-1986 #sequence\_revision 04-Dec-1986 #text\_change 02-Jul-1998

C:Accession: A04017

R:Arya, S.K.; Gallo, R.C.

Proc. Natl. Acad. Sci. U.S.A. 83, 2209-2213, 1986

A:Title: Three novel genes of human T-lymphotropic virus type III: Immune reactivity

A:Reference number: A94093; MUID:86177573

A:Accession: A04017

A:Molecule type: DNA

A:Residues: 1-95 <ARY>

C:Genetics:

A:Gene: tat

C:Superfamily: AIDS trans-activating transcription regulator

C:Keywords: AIDS; immunodeficiency; transcription regulation

Query Match 85.1%; Score 470; DB 1; Length 95;

Best Local Similarity 95.3%; Pred. No. 1.2e-36;

Matches 82; Conservative 2; Mismatches 2; Indels 0; Gaps 0;

Qy 1 MEPVDPRLPEPKHGPSQPKTACTNCYCKKCCFHCQVCFITAAALGISYGRKKRRRRPPQ 60

Db 10 MEPVDPRLPEPKHGPSQPKTACTNCYCKKCCFHCQVCFITAAALGISYGRKKRRRRPPQ 69

Qy 61 GSQTHQVSLSKOPTSQSKGPTGPKE 86

Db 70 GSQTHQVSLSKOPTSQSKGPTGPKE 95

### RESULT 3

A25700

trans-activating transcription regulator - human immunodeficiency virus type 1 (isolate

C:Species: human immunodeficiency virus type 1, HIV-1

C:Date: 28-Sep-1987 #sequence\_revision 28-Sep-1987 #text\_change 02-Jul-1998

C:Accession: A25700

R:Sodroski, J.; Patarca, R.; Rosen, C.; Wong-Staal, F.; Haseltine, W.

Science 229, 74-77, 1985

A:Reference number: A25700; MUID:85244627

A:Accession: A25700

A>Status: preliminary

A:Molecule type: mRNA

A:Residues: 1-86 <SOD>

C:Superfamily: AIDS trans-activating transcription regulator

Query Match 83.7%; Score 462; DB 2; Length 86;

Best Local Similarity 93.0%; Pred. No. 6e-36;

Matches 80; Conservative 2; Mismatches 4; Indels 0; Gaps 0;

Qy 1 MEPVDPRLPEPKHGPSQPKTACTNCYCKKCCFHCQVCFITAAALGISYGRKKRRRRPPQ 60

Db 1 MEPVDPRLPEPKHGPSQPKTACTTCYCKKCCFHCQVCFITAAALGISYGRKKRRRRPPQ 60

Qy 61 GSQTHQVSLSKOPTSQSKGPTGPKE 86

Db 61 GSQTHQVSLSKOPTSQSKGPTGPKE 86

### RESULT 4

T09446

tat protein - human immunodeficiency virus type 1 (strain JRFL)

C:Species: human immunodeficiency virus type 1, HIV-1

C:Date: 16-Jul-1999 #sequence\_revision 16-Jul-1999 #text\_change 26-Aug-1999

C:Accession: T09446

R:Pang, S.; Vinters, H.V.; Akashi, T.; O'Brien, W.A.; Chen, I.S.; Koyanagi, Y.; Namazie,

submitted to the EMBL Data Library, July 1996

A:Reference number: Z16673

A:Accession: T09446

A>Status: preliminary; translated from GB/EMBL/DBJ

A:Molecule type: DNA

A:Residues: 1-101 <PAN>

A:Cross-references: EMBL:U63632; NID:g1465777; PID:g1465783

C:Genetics:

A:Gene: tat

A:Introns: 72/2

C:Superfamily: AIDS trans-activating transcription regulator

Query Match 77.7%; Score 429; DB 2; Length 101;

Best Local Similarity 85.1%; Pred. No. 7.6e-33;

Matches 74; Conservative 4; Mismatches 9; Indels 0; Gaps 0;

Qy 1 MEPVDPRLPEPKHGPSQPKTACTNCYCKKCCFHCQVCFITAAALGISYGRKKRRRRPPQ 60

Db 1 MEPVDPRLPEPKHGPSQPKTACTTCYCKKCCFHCQVCFITAAALGISYGRKKRRRRPPQ 60

Qy 61 GSQTHQVSLSKOPTSQSKGPTGPKE 87

Db 61 GSQTHQVSLSKOPTSQSKGPTGPKE 87

### RESULT 5

E44001

trans-activating transcription regulator - human immunodeficiency virus type 1 (stra

N:Alternate names: tat protein

C:Species: human immunodeficiency virus type 1, HIV-1

A>Note: host Homo sapiens (man)

C:Date: 31-Dec-1993 #sequence\_revision 31-Dec-1993 #text\_change 02-Jul-1998

C:Accession: E44001

R:Li, Y.; Hui, H.; Burgess, C.J.; Price, R.W.; Sharp, P.M.; Hahn, B.H.; Shaw, G.M.

J. Virol. 66, 6587-6600, 1992

A:Title: Complete nucleotide sequence, genome organization, and biological propertie

A:Reference number: A44001; MUID:93021387

A:Accession: E44001

A:Molecule type: DNA

A:Residues: 1-101 <LIY>

A:Cross-references: GB:M93258

C:Genetics:

A:Gene: tat

A:Introns: 72/2

C:Superfamily: AIDS trans-activating transcription regulator

C:Keywords: AIDS; immunodeficiency; transcription regulation

Query Match 77.2%; Score 426; DB 1; Length 101;

Best Local Similarity 83.9%; Pred. No. 1.4e-32;

Matches 73; Conservative 4; Mismatches 10; Indels 0; Gaps 0;

Qy 1 MEPVDPRLPEPKHGPSQPKTACTNCYCKKCCFHCQVCFITAAALGISYGRKKRRRRPPQ 60

Db 1 MEPVDPRLPEPKHGPSQPKTACTTCYCKKCCFHCQVCFITAAALGISYGRKKRRRRPPQ 60

Qy 61 GSQTHQVSLSKOPTSQSKGPTGPKE 87

Db 61 GSQTHQVSLSKOPTSQSKGPTGPKE 87

### RESULT 6

JC5591

transactivator protein - human immunodeficiency virus type 1

N:Alternate names: tat protein

C:Species: human immunodeficiency virus type 1, HIV-1

C:Date: 23-Sep-1997 #sequence\_revision 23-Sep-1997 #text\_change 17-Mar-1999

C:Accession: JC5591

R:Hoffmann, S.; Willbold, D.

Biochem. Biophys. Res. Commun. 235, 806-811, 1997

A:Title: A selection system to study protein-RNA interactions: Functional display of

A:Reference number: JC5591; MUID:97350867

A:Accession: JC5591

A:Molecule type: protein

A:Residues: 1-86 <HO2>

C:Comment: This protein is a key regulatory protein in the viral replication cycle a

C:Superfamily: AIDS trans-activating transcription regulator

F:22-31/Region: cysteine-rich

Query Match 73.0%; Score 403; DB 2; Length 86;

Best Local Similarity 77.9%; Pred. No. 1.7e-30;

Matches 67; Conservative 7; Mismatches 12; Indels 0; Gaps 0;

Qy 1 MEPVDPRLPEPKHGPSQPKTACTNCYCKKCCFHCQVCFITAAALGISYGRKKRRRRPPQ 60

Db 1 MEPVDPRLPEPKHGPSQPKTACTTCYCKKCCFHCQVCFITAAALGISYGRKKRRRRPPQ 60

```
Qy 61 GSQTHQVSLSKOPTSQSKGEPTGPKE 86
      |||||:|||||:|||||:|||||
Db 61 GCQTHQDPPIKQPSQPRGDTGPKE 86

RESULT 7
TNLJZR
trans-activating transcription regulator - human immunodeficiency virus Zr-6
C:Species: human immunodeficiency virus Zr-6
A:Title: Nucleotide sequence of HIV-1-NDK: a highly cytopathic strain of the human imm
C:Date: 30-Sep-1987 #sequence_revision 30-Sep-1987 #text_change 16-Jul-1999
C:Accession: C26192
R:Srinivasan, A.; Anand, R.; York, D.; Ranganathan, P.; Feorino, P.; Schochetman, G.; Cu
Gene 52, 71-82, 1987
A:Title: Molecular characterization of human immunodeficiency virus from Zaire: nucleoti
A:Reference number: A26192; MUID:87248097
A:Accession: C26192
A:Molecule type: DNA
A:Residues: 1-86 <SRI>
A:Cross-references: GB:K03458; GB:M16322; NID:G329398; PIDN:AAA45377.1; PID:G329400
C:Genetics:
A:Gene: tat
A:Introns: 72/3
C:Superfamily: AIDS trans-activating transcription regulator
C:Keywords: AIDS; immunodeficiency; transcription regulation

Query Match 72.5%; Score 400; DB 1; Length 86;
Best Local Similarity 77.9%; Pred. No. 3.2e-30;
Matches 67; Conservative 6; Mismatches 13; Indels 0; Gaps 0;

Qy 1 MEPVDPRLEPKHPSQPKTACTNCYCKKCCFHCQVCFITAAALGISYGRKKRRRRPPQ 60
      |||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
Db 1 MDPVDPNLEPNHPSQPKTACNCRCKKCCYHCQVCFITKGLGISYGRKKRRRRPSQ 60

Qy 61 GSQTHQVSLSKOPTSQSKGEPTGPKE 86
      |||||:|||||:|||||:|||||
Db 61 GCQTHQDPPIKQPSQPRGDTGPKE 86

RESULT 8
S54381
tat protein - human immunodeficiency virus type 1
C:Species: human immunodeficiency virus type 1, HIV-1
C:Date: 15-Jul-1995 #sequence_revision 01-Sep-1995 #text_change 20-Sep-1999
C:Accession: S54381
R:Theodore, T.; Buckler-White, A.J.
Submitted to the EMBL Data Library, July 1989
A:Reference number: S54377
A:Accession: S54381
A:Status: preliminary
A:Molecule type: genomic RNA
A:Residues: 1-86 <THE>
A:Cross-references: EMBL:M2639; NID:G329377; PIDN:AAA45363.1; PID:G329378
C:Genetics:
A:Introns: 72/2
C:Superfamily: AIDS trans-activating transcription regulator

Query Match 72.5%; Score 400; DB 2; Length 86;
Best Local Similarity 76.7%; Pred. No. 3.2e-30;
Matches 66; Conservative 8; Mismatches 12; Indels 0; Gaps 0;

Qy 1 MEPVDPRLEPKHPSQPKTACTNCYCKKCCFHCQVCFITAAALGISYGRKKRRRRPPQ 60
      |||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
Db 1 MDPVDPNLEPNHPSQPKTACNCRCKKCCYHCQVCFITKGLGISYGRKKRRRRPSQ 60

Qy 61 GSQTHQVSLSKOPTSQSKGEPTGPKE 86
      |||||:|||||:|||||:|||||
Db 61 GCQTHQDPPIKQPSQPRGDTGPKE 86

RESULT 9
trans-activating transcription regulator - human immunodeficiency virus type 1 (isola
C:Species: human immunodeficiency virus type 1, HIV-1
C:Date: 30-Jun-1988 #sequence_revision 30-Jun-1988 #text_change 02-Jul-1998
C:Accession: B25523
R:Desai, S.M.; Kalyanaraman, V.S.; Casey, J.M.; Srinivasan, A.; Andersen, P.R.; Devar
```

```
TNLJND
trans-activating transcription regulator - human immunodeficiency virus type 1 (isola
C:Species: human immunodeficiency virus type 1, HIV-1
A:Date: 30-Sep-1991 #sequence_revision 30-Sep-1991 #text_change 16-Jul-1999
C:Accession: JQ0071
R:Spire, B.; Sire, J.; Zachar, V.; Rey, F.; Barre-Sinoussi, F.; Galibert, F.; Hampe,
Gene 81, 275-284, 1989
A:Title: Nucleotide sequence of HIV1-NDK: a highly cytopathic strain of the human imm
A:Reference number: JQ0065; MUID:90034200
A:Accession: JQ0071
A:Molecule type: DNA
A:Residues: 1-86 <SPT>
A:Cross-references: GB:M27323; NID:G328154; PIDN:AAA44866.1; PID:G328155
C:Genetics:
A:Gene: tat
A:Superfamily: AIDS trans-activating transcription regulator
C:Keywords: AIDS; immunodeficiency; transcription

Query Match 71.6%; Score 395; DB 1; Length 86;
Best Local Similarity 74.4%; Pred. No. 9.1e-30;
Matches 64; Conservative 11; Mismatches 11; Indels 0; Gaps 0;

Qy 1 MEPVDPRLEPKHPSQPKTACTNCYCKKCCFHCQVCFITAAALGISYGRKKRRRRPPQ 60
      |||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
Db 1 MDPVDPNLEPNHPSQPKTACNCRCKKCCYHCQVCFITKGLGISYGRKKRRRRPPQ 60

Qy 61 GSQTHQVSLSKOPTSQSKGEPTGPKE 86
      |||||:|||||:|||||:|||||
Db 61 GDOAHQVPIPEQPSQSGDPTGPKE 86

RESULT 10
T01665
tat protein - human immunodeficiency virus type 1
C:Species: human immunodeficiency virus type 1, HIV-1
C:Date: 19-Feb-1999 #sequence_revision 19-Feb-1999 #text_change 21-Jul-2000
C:Accession: T01665
R:Alizon, M.; Wain-Hobson, S.; Gluckman, J.C.; Sonigo, P.
Cell 46, 63-74, 1986
A:Title: Genetic variability of the AIDS virus: Nucleotide sequence analysis of two 1
A:Reference number: 214389; MUID:86245056
A:Accession: T01665
A:Status: preliminary; translated from GB/EMBL/DBJ
A:Molecule type: mRNA
A:Residues: 1-87 <ALI>
A:Cross-references: EMBL:K03456; NID:G60228; PIDN:CAA28015.1; PID:G60233
C:Genetics:
A:Introns: 72/2
C:Superfamily: AIDS trans-activating transcription regulator

Query Match 70.2%; Score 387.5; DB 2; Length 87;
Best Local Similarity 74.7%; Pred. No. 4.5e-29;
Matches 65; Conservative 9; Mismatches 12; Indels 1; Gaps 1;

Qy 1 MEPVDPRLEPKHPSQPKTACTNCYCKKCCFHCQVCFITAAALGISYGRKKRRRRPPQ 60
      |||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
Db 1 MDPVDPNLEPNHPSQPKTACNCRCKKCCYHCQVCFITKGLGISYGRKKRRRRPPQ 60

Qy 61 GSQTHQVSLSKOPTSQSKGEPTGPKE 86
      |||||:|||||:|||||:|||||
Db 61 GNOAHQDPLPEQPSQHRGDTGPKE 87

RESULT 11
TNLJH4
trans-activating transcription regulator - human immunodeficiency virus type 1 (isola
C:Species: human immunodeficiency virus type 1, HIV-1
C:Date: 30-Jun-1988 #sequence_revision 30-Jun-1988 #text_change 02-Jul-1998
C:Accession: B25523
R:Desai, S.M.; Kalyanaraman, V.S.; Casey, J.M.; Srinivasan, A.; Andersen, P.R.; Devar
```

PROC. Natl. Acad. Sci. U.S.A. 83, 8380-8384, 1986  
 A:Title: Molecular cloning and primary nucleotide sequence analysis of a distinct human  
 A:Reference number: A94136; MUID:87041461  
 A:Accession: B25523  
 A:Molecule type: DNA  
 A:Residues: 1-72 <DES>  
 A:Cross-references: GB:M13137; NID:g326460  
 A:Note: the GenBank entry ADRE3AA PID:g209908 differs from the published sequence in tra  
 C:Gene: tat  
 C:Superfamily: AIDS trans-activating transcription regulator  
 C:Keywords: transcription regulation

Query Match 65.9%; Score 364; DB 1; Length 72;  
 Best Local Similarity 88.9%; Pred. No. 5.7e-27;  
 Matches 64; Conservative 0; Mismatches 8; Indels 0; Gaps 0;  
 Qy 1 MEPVDPRLPEPKHGPSQPKTACTNCYCKKCCFHCOVCFITAAALGISYGRKKRRRPPQ 60  
 Db 1 MEPVDPRLPEPKHGPSQPKTACTNCYCKKCCFHCOVCFITAAALGISYGRKKRRRRAHQ 60  
 Qy 61 GSQTHQVLSKQ 72  
 Db 61 DSONHOASLSKQ 72

RESULT 12  
 T09384  
 trans-activating transcription regulator - human immunodeficiency virus type 1 (isolate  
 C:Species: human immunodeficiency virus type 1, HIV-1  
 C:Date: 11-Jun-1999 #sequence\_revision 11-Jun-1999 #text\_change 21-Jul-2000  
 C:Accession: T09384  
 R:Michael, N.L.; Chang, G.; d'ARCY, L.A.; Ehrenberg, P.K.; Mariani, R.; Busch, M.P.; Bid  
 J. Virol. 69, 4228-4236, 1995  
 A:Title: Defective accessory genes in a human immunodeficiency virus type 1-infected lon  
 A:Reference number: 216654; MUID:95287475  
 A:Accession: T09384  
 A>Status: preliminary; translated from GB/EMBL/DBJ  
 A:Molecule type: DNA  
 A:Residues: 1-71 <MC>  
 A:Cross-references: EMBL:U24451; NID:9829440; PIDN:AAA79576.1; PID:g829444  
 C:Gene: tat  
 C:Superfamily: AIDS trans-activating transcription regulator  
 C:Keywords: transcription

Query Match 63.9%; Score 353; DB 2; Length 71;  
 Best Local Similarity 83.1%; Pred. No. 5.9e-26;  
 Matches 59; Conservative 4; Mismatches 8; Indels 0; Gaps 0;  
 Qy 1 MEPVDPRLPEPKHGPSQPKTACTNCYCKKCCFHCOVCFITAAALGISYGRKKRRRPPQ 60  
 Db 1 MEPVDPRLPEPKHGPSQPKTACNCCYCKKCCFHCOVCFVKGKLGISYGRKKRRRSPQ 60  
 Qy 61 GSQTHQVLSK 71  
 Db 61 DSEAHQASLSE 71

RESULT 13  
 TNLJST  
 trans-activating transcription regulator - simian immunodeficiency virus SIVcpz  
 C:Species: simian immunodeficiency virus SIVcpz  
 A:Note: host Pan troglodytes (chimpanzee)  
 C:Date: 30-Sep-1991 #sequence\_revision 30-Sep-1991 #text\_change 16-Jul-1999  
 C:Accession: S09987  
 R:Huet, T.; Cheynier, R.; Meyerhans, A.; Roelants, G.; Wain-Hobson, S.  
 Nature 345, 356-359, 1990  
 A:Title: Genetic organization of a chimpanzee lentivirus related to HIV-1.  
 A:Reference number: S09983; MUID:90259077  
 A:Accession: S09987

A>Status: nucleic acid sequence not shown; translation not shown  
 A:Molecule type: DNA  
 A:Residues: 1-100 <HUE>  
 A:Cross-references: EMBL:X52154; NID:g58866; PIDN:CAA36404.1; PID:g763084  
 C:Gene: tat  
 A:Introns: 73/2  
 A:Superfamily: AIDS trans-activating transcription regulator  
 C:Keywords: AIDS; immunodeficiency; transcription

Query Match 57.2%; Score 315.5; DB 1; Length 100;  
 Best Local Similarity 60.7%; Pred. No. 2.2e-22;  
 Matches 54; Conservative 12; Mismatches 22; Indels 1; Gaps 1;  
 Qy 1 MEPVDPRLPEPKHGPSQPKTACTNCYCKKCCFHCOVCFITAAALGISYGRKKRRRPP 59  
 Db 1 MDPIDPDLPEPKHGPSQPKTVCNCCYCKKCCYHCYCFCTKKGLGISYGRKKRTTTRTAP 60  
 Qy 60 QGSOHQVLSKQTSOSKGEPTGPKETS 88  
 Db 61 AGSKNODSIPKQPLSOSRGNKEGSEKST 89

RESULT 14  
 TNLJGG  
 trans-activating transcription regulator - human immunodeficiency virus type 2 (isol:  
 N:Alternate names: tat protein  
 C:Species: human immunodeficiency virus type 2, HIV-2  
 A:Note: host Homo sapiens (man)  
 C:Date: 30-Jun-1990 #sequence\_revision 30-Jun-1990 #text\_change 16-Jul-1999  
 C:Accession: J50332  
 R:Hasegawa, A.; Tsujimoto, H.; Maki, N.; Ishikawa, K.; Miura, T.; Fukasawa, M.; Miki  
 AIDS Res. Hum. Retroviruses 5, 593-604, 1989  
 A:Title: Sequence of a distinct HIV-2 isolate from Ghana showing significant diverge  
 A:Reference number: J50327; MUID:90122350  
 A:Accession: J50332  
 A:Molecule type: DNA  
 A:Residues: 1-130 <HAS>  
 A:Cross-references: GB:M30895; GB:D00477; NID:g325709; PIDN:AAA3929.1; PID:g325710  
 A:Note: This sequence was submitted to JIPID, October 1989  
 C:Gene: tat  
 A:Gene: tat  
 A:Introns: 99/2  
 C:Superfamily: AIDS trans-activating transcription regulator  
 C:Keywords: AIDS; transcription regulation

Query Match 31.8%; Score 175.5; DB 1; Length 130;  
 Best Local Similarity 49.3%; Pred. No. 2.2e-09;  
 Matches 35; Conservative 9; Mismatches 24; Indels 3; Gaps 2;  
 Qy 17 QPKTACTN-CYCKKCCFHCOVCFITAAALGISYGRKKRRRPPQSGSOHQVLSKQPTS 75  
 Db 45 RPLEACTNSCYCKQSPFCFLKGLGIWYAKSR--RRTPRKTKTHSSASDKSIS 102  
 Qy 76 QSKGEPTGPK 86  
 Db 103 TRTGDSDQPTKE 113

RESULT 15  
 TNLJST  
 trans-activating transcription regulator - human immunodeficiency virus type 2 (isol:  
 N:Alternate names: tat protein  
 C:Species: human immunodeficiency virus type 2, HIV-2  
 A:Note: host Homo sapiens (man)  
 C:Date: 31-Dec-1990 #sequence\_revision 31-Dec-1990 #text\_change 16-Jul-1999  
 C:Accession: F33943  
 R:Kumar, P.; Hui, H.; Kappes, J.C.; Haggarty, B.S.; Hoxie, J.A.; Arya, S.K.; Shaw, G  
 J. Virol. 64, 890-901, 1990  
 A:Title: Molecular characterization of an attenuated human immunodeficiency virus tvi  
 A:Reference number: A33943; MUID:90112662

```
A:Gene: tat
A:Introns: 99/3
C:Superfamily: AIDS trans-activating transcription regulator
C:Keywords: AIDS; immunodeficiency; transcription regulation

Query Match          29.6%; Score 163.5; DB 1; Length 130;
Best Local Similarity 42.7%; Pred. No. 2.8e-08;
Matches 35; Conservative 10; Mismatches 24; Indels 13; Gaps 4;

QY   17  QPKTACTN-CYCKKCFHCVOCPITAAIGISYGRKKRRRRPPQGSGQTH-----QVSLS 70
      :|::||::||::||::||::||::||::||::||::||::||::||::||::||::||:
Db    45  RPLETCNNCYCKRCVCVHQCMLFGLGICYERKGR--RRRTPKKTHTPSPDPKSIS 102

QY   71  -----KQPTSQSKGEPTGPKET 87
      |||::|
Db    103 TRTGDSQPTTKQRKVTAEVTET 124

RESULT 18
TNLYCA
trans-activating transcription regulator - human immunodeficiency virus type 2 (isla
N:Alternate names: tat protein
C:Species: human immunodeficiency virus type 2, HIV-2
A>Note: host Homo sapiens (man)
C>Date: 31-Mar-1992 #sequence_revision 31-Mar-1992 #text_change 16-Jun-2000
C:Accession: I38475; JQ0981
R:Fristem, M.; Hill, F.; Karpas, A.
J. Gen. Virol. 72, 721-724, 1991
A>Title: Nucleotide sequence of a Guinea-Bissau-derived human immunodeficiency virus
A:Reference number: A38475; MUID:91170959
A:Accession: I38475
A:Molecule type: DNA
A:Residues: 1-133 <TRI>
A:Cross-references: GB:D00835; NID:g3153166; PIDN:BAA00714.1; PID:g221467
C:Genetics:
A:Gene: tat
A:Introns: 102/2
C:Superfamily: AIDS trans-activating transcription regulator
C:Keywords: transcription
```

```

Query Match          29.3%; Score 161.5; DB 1; Length 133;
Best Local Similarity 43.8%; Pred. No. 4.4e-08;
Matches 32; Conservative 8; Mismatches 20; Indels 13; Gaps 3;

Qy 17 QPKTACTN-CYCKKCCFHQCVCFTAAALGTSYGRKKRRQRRPPQGSQTHQVSL----- 70
      ||| ||||| ||||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 48 QPLEECDNSCYKRCQCYHCQLCFLLKGLGICYDEKGR--RRTPPKAKAHSSASDKSIS 105
      ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||

Qy 71 -----KQPTQSOK 78
      || : |
Db 106 TRTRNSQPAKKQK 118
      || : |

RESULT 19
SI2157
trans-activating transcription regulator - human immunodeficiency virus type 2
C:Species: human immunodeficiency virus type 2, HIV-2
C:Date: 21-Nov-1993 #sequence_revision 10-Nov-1995 #text_change 20-Sep-1999
C:Accession: SI2157
R:Kuehnelt, H.; Kreutz, R.; Ruebsamen-Waigmann, H.
Nucleic Acids Res. 19, 6142, 1990
A:Title: Nucleotide sequence of HIV-2(D194), an isolate from a Gambian case of 'Neuro
A:Reference number: SI2152; MOID:91045094
A:Accession: SI2157
A:Status: preliminary; nucleic acid sequence not shown; translation not shown
A:Molecule type: DNA
A:Residues: 1-130 <KUE>
A:Cross-references: EMBL:X52223; NID:g60155; PIDN:CAN36459.1; PID:g763106
A:Note: the nucleotide sequence was submitted to the EMBL Data Library, June 1990
C:Genetics:
A:Introns: 99/2

```

C:Superfamily: AIDS trans-activating transcription regulator  
C:Keywords: transcription regulation

Query Match 29.2%; Score 161; DB 2; Length 130;  
Best Local Similarity 33.9%; Pred. No. 4.8e-08;  
Matches 39; Conservative 12; Mismatches 26; Indels 38; Gaps 5;  
QY 2 EPVDRPRLPKHPSQ-----PKTACTN-CYCKKCCFHQCV 36  
DB 7 EP-ESSLESYNPSSCTSERDVTQAKRQGBELLAQLHRPLEACTNSCYKQCSYHCQL 65  
QY 37 CFITAAALGISYCRKRRORRRPPQGSQTHQVSL 81  
DB 66 CFLKGLGIWYARQGR--RRRTPRKTKTHPPASDKSISTRGDSOPTKKQKTP 118

RESULT 20  
TNLJG3  
trans-activating transcription regulator - simian immunodeficiency virus (macaque isolate)  
C:Species: simian immunodeficiency virus, SIV  
C>Date: 30-Jun-1989 #sequence\_revision 30-Jun-1989 #text\_change 02-Jul-1998  
C:Accession: F28887  
R:Chakrabarti, L.; Guyader, M.; Allizon, M.; Daniel, M.D.; Desrosiers, R.C.; Tiollais, P.  
Nature 328, 543-547, 1987  
A>Title: Sequence of simian immunodeficiency virus from macaque and its relationship to  
A:Reference number: A28887; MUID:87287230  
A:Molecule type: DNA  
A:Residues: 1-129 <CHA>  
A:Cross-references: GB:Y00277; GB:M16403; NID:g61730  
C:Genetics: tat  
A:Gene: tat  
A:Introns: 98/3  
C:Superfamily: AIDS trans-activating transcription regulator  
C:Keywords: AIDS; immunodeficiency; transcription

Query Match 27.4%; Score 151; DB 1; Length 129;  
Best Local Similarity 43.1%; Pred. No. 4e-07;  
Matches 31; Conservative 9; Mismatches 20; Indels 12; Gaps 3;  
QY 17 QPKTACTN-CYCKKCCFHQCVCFITAAALGISYGRKKRRORRRPPQGSQTHQVLSK---- 71  
DB 45 RPLEACNTCYCKKCCYHCQFCFLKGLGISY--EKSHRRRTPRKAKANTSSASNPIN 102  
QY 72 -----QPTSQK 78  
DB 103 RIRLCQPKKAKK 114

RESULT 21  
TNLJS2  
trans-activating transcription regulator - simian immunodeficiency virus SIVagm (type 3)  
C:Species: simian immunodeficiency virus SIVagm  
C>Date: 31-Mar-1989 #sequence\_revision 31-Mar-1989 #text\_change 02-Jul-1998  
C:Accession: A26737  
R:Hirsch, V.; Riedel, N.; Mullins, J.I.  
Cell 49, 307-319, 1987  
A>Title: The genome organization of STLV-3 is similar to that of the AIDS virus except f  
A:Reference number: A26737; MUID:87187627  
A:Accession: A26737  
A:Molecule type: DNA  
A:Residues: 1-106 <HIR>  
A:Cross-references: GB:M19499; NID:g334657  
A>Note: the authors translated the codon GAC for residue 9 as Asn  
C:Genetics: tat  
A:Gene: tat  
A:Introns: 74/2  
C:Superfamily: AIDS trans-activating transcription regulator  
C:Keywords: AIDS; transcription

Query Match 27.0%; Score 149; DB 1; Length 106;  
Best Local Similarity 43.2%; Pred. No. 5.3e-07;  
Matches 32; Conservative 13; Mismatches 23; Indels 6; Gaps 3;

QY 1 MEVDPRLPKHPSQ---PKTACTN-CYCKKCCFHQCVCFITAAALGISYGRKKRRORRR 56  
DB 1 MOPLONRDPLGEEILSYLRPLEACNTCYCKKCCYHCQFCFLKGLGICY--EOSRRR 58  
QY 57 RPPQGSQTHQVSL 70  
DB 59 RTPKKAKANTSSAS 72

RESULT 22  
S08440  
trans-activating transcription regulator - human immunodeficiency virus type 2 D205  
C:Species: human immunodeficiency virus type 2 D205, HIV-2 D205  
C>Date: 07-Sep-1990 #sequence\_revision 07-Sep-1990 #text\_change 02-Jul-1998  
C:Accession: S08440  
R:Dietrich, U.; Adamski, M.; Kreutz, R.; Seipp, A.; Kuehn, H.; Ruebsamen-Waigmann,  
Nature 342, 948-950, 1989  
A>Title: A highly divergent HIV-2-related isolate.  
A:Reference number: S08434; MUID:90081881  
A:Accession: S08440  
A>Status: nucleic acid sequence not shown; translation not shown  
A:Molecule type: DNA  
A:Residues: 1-96 <DIE>  
A:Cross-references: EMBL:X16109  
A>Note: this sequence was submitted to the EMBL Data Library, Aug-1989  
C:Genetics: tat  
A:Gene: tat  
C:Superfamily: AIDS trans-activating transcription regulator  
C:Keywords: transcription regulation

Query Match 26.9%; Score 148.5; DB 2; Length 96;  
Best Local Similarity 57.1%; Pred. No. 5.4e-07;  
Matches 24; Conservative 6; Mismatches 11; Indels 1; Gaps 1;  
QY 17 QPKTACTN-CYCKKCCFHQCVCFITAAALGISYGRKKRRORRR 57  
DB 45 RPLKACDNTCYCKKCCYHCQFCFLKGLGICYDRSRRSRAKR 86

RESULT 23  
TNLJG4  
trans-activating transcription regulator - simian immunodeficiency virus (African pr  
C:Species: simian immunodeficiency virus, SIV  
C>Date: 30-Jun-1989 #sequence\_revision 30-Jun-1989 #text\_change 28-Jul-2000  
C:Accession: E30045  
R:Fukasawa, M.; Miura, T.; Hasegawa, A.; Morikawa, S.; Tsujimoto, H.; Miki, K.; Kita  
Nature 333, 457-461, 1988  
A>Title: Sequence of simian immunodeficiency virus from African green monkey, a new  
A:Reference number: A30045; MUID:88232906  
A:Accession: E30045  
A:Molecule type: DNA  
A:Residues: 1-100 <FUK>  
A:Cross-references: EMBL:X07805; NID:g61748; PIDN:CAA30661.1; PID:g4469309  
C:Genetics: tat  
A:Gene: tat  
A:Introns: 73/2  
C:Superfamily: AIDS trans-activating transcription regulator  
C:Keywords: AIDS; immunodeficiency; transcription

Query Match 25.7%; Score 142; DB 1; Length 100;  
Best Local Similarity 39.5%; Pred. No. 2.2e-06;  
Matches 30; Conservative 9; Mismatches 27; Indels 10; Gaps 3;  
QY 17 QPKTACTN-CYCKKCCFHQCVCFITAAALGISY-----GRKKRRORRRPPQGSQTHQVSL 70  
DB 21 KPLQTKCKKCCYHCQFCFLKGLGIVTHAPTRRKRKRLSLNLAFLQHSISTKMG 80



```

Qy      71 KQ - - - - P T S Q S K G E P T      82
      :      | | | | | | | |
Db      81 R D G O T T P T S O E K V E T T      96

```

RESULT 24

A46356  
tat protein - simian immunodeficiency virus SIVagm (type 3)  
C:Species: simian immunodeficiency virus SIVagm  
C:date: 03-Feb-1994 #sequence\_revision 03-Feb-1994 #text\_change 20-Sep-1999  
C:Accession: A46356; A36860  
R:Baier, M.; Garber, C.; Mueller, C.; Cichutek, K.; Kurth, R.  
Virology 176, 216-221, 1990  
A:Title: Complete nucleotide sequence of a simian immunodeficiency virus from African gray monkey  
A:Reference number: A46356; MUID:90232731  
A:Accession: A46356  
A:Status: nucleic acid sequence not shown  
A:Molecule type: DNA  
A:Residues: 1-119 <BAI>  
A:Cross-references: GB:M30931; NID:G334400; PID:AAA91917.1; PID:G334401  
C:Genetics:  
A:Introns: 73/2  
C:Superfamily: AIDS trans-activating transcription regulator  
C:Keywords: transcription regulation

Query Match	25.3%	Score 139.5;	DB 2;	Length 119;
Best Local Similarity	30.6%;	Pred. No. 4.3e-06;		
Matches 30;	Conservative	14;	Mismatches 25;	Indels 29;
Gaps				

Qy 18 PKTACTN-CYCKKCCFHCQVCFITAAALGISYGRKKRRQRPP-----QG 61  
| | | | | | | | | | | | | | | | : : :  
Db 22 PLKRGTKNCYCKCCCCHCQLCFLOKGLGVTHAPRIIRKKLAPLDRFPEOKQSISTRGRD 81

62 SOTHO-----VSLSKOPTSOSKGEPTGPKE 86

82 SQTQKGQEKVETSARTAPSLGRKNLAQQSGRATGASD 119

RESULT 25

T11564 25  
tat protein - simian immunodeficiency virus SIVsm (strain E543)  
C:Species: simian immunodeficiency virus SIVsm  
A:Variety: strain E543  
C:Date: 16-Jul-1999 #sequence\_revision 16-Jul-1999 #text\_change 21-Jul-2000  
C:Accession: T11564  
R:Hirsch, V.M.; Adger-Johnson, D.; Cambell, B.; Goldstein, S.; Brown, C.; Elkins, W.R.; J. Virol. 71, 1608-1620, 1997  
A:Title: A molecularly cloned, pathogenic, neutralization-resistant simian immunodeficiency virus  
A:Reference number: Z17285; MUID:97151152  
A:Accession: T11564  
A:Status: preliminary; translated from GE/EMBL/DBDJ  
A:Molecule type: DNA  
A:Residues: 1-132 <HIR>  
A:Cross-references: EMBL:U72748; NID:g1695908; PIDN:AAC56563.1; PID:g1695914  
C:Genetics:  
A:Gene: tat  
A:Introns: 100/2  
C:Superfamily: AIDS trans-activating transcription regulator  
C:Keywords: AIDS; immunodeficiency

Query Match	25.1%	Score 138.5;	DB 2;	Length 132;
Best Local Similarity	36.6%	Pred. NO. 5.8e-06;		
Matches 26;	Conservative 14;	Mismatches 26;	Indels 5;	Gaps 2;

Qy	17	QPKTACTN-CYCKKCCFHCQVCFITAAALGISYGRKKRRQRPPQGSOHQVLSLKQPTS	75
	:	:         :         :         :         :	:
Db	48	RPLEPCYNKYCRKCCHCEFLKKGIGICY----EHHRRTPKTKTNPLPASNNRSL	103

Qv 76 0SKCEPDTGPKF 86

WY 70 Q3KGEFTGFK 60  
:: ||:

Db 104 STRTRNRQPKK 114

**RESULT** 26

T11557  
tat protein - simian immunodeficiency virus SIVsm  
C:Species: simian immunodeficiency virus SIVsm  
A:Variety: strain 62  
C>Date: 16-Jul-1999 #sequence\_revision 16-Jul-1999 #text\_change 26-Aug-1999  
R/Accession: T11557  
R/Hirsch, V.M.; Martin, J.E.; Dapolito, G.; Elkins, W.R.; London, W.T.; Goldstein, S.  
J. Virol. 68, 2649-2661, 1994  
A:Title: Spontaneous substitutions in the vicinity of the V3 analog affect cell tropi  
A:Reference number: 217284; MUID:94187106  
A:Accession: T11557  
A:Status: preliminary; translated from GB/EMBL/DBDJ  
A:Molecule type: DNA  
A:Residues: 1-131 <HIR>  
A:Cross-references: EMBL:U04985; NID:g451609; PID:g451613  
C:Genetics:  
A:Gene: tat  
A:Introns: 100/2  
C:Superfamily: AIDS trans-activating transcription regulator  
C:Keywords: AIDS; immunodeficiency

Query Match	24.2%	Score 133.5;	DB 2;	Length 131;
Best Local Similarity	38.4%	Pred. No. 1.7e-05;		
Matches 28;	Conservative	12;	Mismatches 18;	Indels 15;
				Gaps 4;

QY 17 QPKTACTN-CYCKKCCFHCQCVCFITAAIGISYGRKKRRQRPPQGSOTHQV-----SLS 70  
:  
: | | | | | : | | | | : | | | | : | | | | : | | | | : | | | | : | | | | : | | | | :  
Db 48 RPLETCYNKYCKRCYCHCQHGLKKGIGICY-----EOHRRTPKTKKANPLAPASNKSL 10

0v 71 -----K0PTC0CK 78

QY	71	KQF13Q3K	78
		:   :	
Dp	104	TRARNROPKKEK	116

RESULT 27

A48344  
trans-activating transcription regulator - simian immunodeficiency virus (isolate Afr  
C:Species: simian immunodeficiency virus, SIV  
C:Date: 17-Feb-1994 #sequence\_revision 17-Feb-1994 #text\_change 02-Jul-1998  
C:Accession: A48344  
R:Sakai, H.; Sakuragi, J.; Sakuragi, S.; Shibata, R.; Hayami, M.; Ishimoto, A.; Adach  
Arch. Virol. 125, 1-14, 1992  
A:Title: Genetic characterization of simian immunodeficiency virus isolated from an A  
A:Reference number: A48344; MUID:92352315  
A:Accession: A48344  
A:Molecule type: DNA  
A:Residues: 1-116 <SAK>  
A:Note: sequence extracted from NCBI backbone (NCBIP:109904)  
C:Genetics:  
A:Gene: tat  
C:Superfamily: AIDS trans-activating transcription regulator  
C:Keywords: AIDS; immunodeficiency; transcription

Query Match	22.3%	Score 123;	DB 1;	Length 116;
Best Local Similarity	44.2%	Pred. NO. 0.00014;		
Matches 23:	Conservative	8:	Mismatches 17:	Indels 4:
				Gaps 2:

**QY** 16 SQPKTACTN-CYCKKCCFHCQVCFITAAIGISYGRKKRR---QRRRPDQSQ 63  
:  
:  
**Dd** 31 TOPLOACENKCWCCKCFHCLMFCFKKKGIGIRYHVYRKRGPSTNKKIPGGE 82  
:

SECRET

trans-activating transcription regulator - bovine immunodeficiency virus (isolate 127  
TNLJBT  
C:Species: bovine immunodeficiency virus  
C:Date: 30-Sep-1991 #sequence\_revision 30-Sep-1991 #text\_change 02-Jul-1998  
C:Accession: D34742

R:Garvey, K.J.; Oberste, M.S.; Else, J.E.; Braun, M.J.; Gonda, M.A.  
Virology 175, 391-409, 1990

A:Title: Nucleotide sequence and genome organization of biologically active proviruses of  
A:Reference number: A34742; MUID:90223985

A:Accession: D34742

A:Molecule type: genomic RNA

A:Residues: 1-145 <GAR>

A:Cross-references: GB:M32690

C:Genetics:

A:Gene: tat

A:Introns: 103/3

C:Superfamily: AIDS trans-activating transcription regulator

C:Keywords: AIDS; immunodeficiency; transcription

Query Match 20.7%; Score 114.5; DB 1; Length 145;  
Best Local Similarity 32.9%; Pred. No. 0.001;  
Matches 23; Conservative 10; Mismatches 28; Indels 9; Gaps 1;

Qy 18 PKTACTNCCYCKKCCFHCVCFTTAALGISYGR-----KKRQRRRPPQGSOTHOVS 68

Db 34 PRDCPHCCPCISWHCOLCEFLQNLGINYGSPRPRTRGKGRIRRTASGGDQREAD 93

Qy 69 LSKOPTSQSK 78

Db 94 SORSFTNMDQ 103

RESULT 29

S46349

trans-activating transcription regulator - simian immunodeficiency virus SIVagm (isolate  
N:Alternate names: tat protein

C:Species: simian immunodeficiency virus SIVagm

A:Variety: isolate SAB-1

C:Date: 25-Dec-1994 #sequence\_revision 14-Feb-1997 #text\_change 20-Sep-1999

C:Accession: S46349

R:Jin, M.J.; Hui, H.; Robertson, D.L.; Mueller, M.C.; Barre-Sinoussi, F.; Hirsch, V.M.;

EMBO J. 13, 2935-2947, 1994

A:Title: Mosaic genome structure of simian immunodeficiency virus from West African green  
A:Reference number: S46349; MUID:94298785

A:Accession: S46349

A>Status: nucleic acid sequence not shown; translation not shown

A:Molecule type: DNA

A:Residues: 1-73 <JIN>

A:Cross-references: EMBL:U04005; NID:9466229; PIDN:AAA21507.1; PID:9466233

A:Experimental source: isolate SAB-1; sabaeus monkey

A:Note: the nucleotide sequence was submitted to the EMBL Data Library, December 1993

C:Genetics:

A:Gene: tat

C:Superfamily: AIDS trans-activating transcription regulator

Query Match 20.3%; Score 112; DB 2; Length 73;

Best Local Similarity 46.3%; Pred. No. 0.001;

Matches 25; Conservative 6; Mismatches 19; Indels 4; Gaps 2;

Qy 17 OPTACTN-CYCKKCCFHCVCFTTAALGISYGRKKRQRRRPPQGSOTHOVSL 69

Db 21 RPLQACDNTCFCKVCCFHCILCFHKKALGIRYVPR---PRRASKKISHNOVSL 71

RESULT 30

T28897

hypothetical protein T17H7.5 - Caenorhabditis elegans

C:Species: Caenorhabditis elegans

C:Date: 29-Oct-1999 #sequence\_revision 29-Oct-1999 #text\_change 29-Oct-1999

C:Accession: T28897

R:Favella, A.

submitted to the EMBL Data Library, January 1996

A:Description: The sequence of C. elegans cosmid T17H7.

A:Reference number: Z20340

A:Accession: T28897

A>Status: preliminary; translated from GB/EMBL/DBJ

A:Molecule type: DNA

A:Residues: 1-1001 <FAV>

A:Cross-references: EMBL:U42841; PIDN:AA48171.1; GSPDB:GN00021; CESP:T17H7.5

A:Experimental source: strain Bristol N2; clone T17H7

C:Genetics:

A:Gene: CESP:T17H7.5

A:Map position: 3

A:Introns: 50/3; 140/3; 762/3; 795/3; 840/3; 942/3

Query Match 15.9%;

Best Local Similarity 38.2%; Score 88; DB 2; Length 1001;

Matches 21; Conservative 5; Mismatches 15; Indels 14; Gaps 2;

Qy 42 ALGISYGRKKRRRPPQGSQ--THQVLSKQPTSQSKGEPTGPKETSGHHHHH 95

Db 856 SLMPHHRQSRSDSRGPPQMSRAASSIFLSPQTPQ-----HHHHHH 897

Search completed: August 26, 2002, 08:13:24

Job time: 293 sec

GenCore version 4.5  
Copyright (c) 1993 - 2000 Compugen Ltd.

OM protein - protein search, using sw model

Run on: August 26, 2002, 08:15:43 ; Search time 40.39 Seconds  
(without alignments)  
91.071 Million cell updates/sec

Title: US-09-509-239-23  
Perfect score: 552  
Sequence: 1 MEPVDPRLPKWPKSQPKT.....QSKGPTGPKTSGHHHHH 95

Scoring table: BLOSUM62  
Gapop 10.0 , Gapext 0.5  
Searched: 105224 seqs, 38719550 residues

Total number of hits satisfying chosen parameters: 105224

Minimum DB seq length: 0  
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

Database : SwissProt\_40:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

## SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	478	86.6	86	1 TAT_HV1B1	P04606 human immun
2	473	85.7	86	1 TAT_HV1P1	P04607 human immun
3	470	85.1	86	1 TAT_HV112	P04326 human immun
4	462	83.7	86	1 TAT_HV1B1R	P04610 human immun
5	453	82.1	102	1 TAT_HV1RH	P05908 human immun
6	446	80.8	86	1 TAT_HV1H2	P04608 human immun
7	438	79.3	101	1 TAT_HV1JR	P08079 human immun
8	430	77.9	101	1 TAT_HV1S1	P19553 human immun
9	429	77.7	101	1 TAT_HV1SC	P05906 human immun
10	427	77.4	101	1 TAT_HV1C4	P05907 human immun
11	426	77.2	101	1 TAT_HV1Y2	P35965 human immun
12	423	76.6	101	1 TAT_HV1MN	P05905 human immun
13	423	76.6	101	1 TAT_HV1S3	P19552 human immun
14	411	74.5	101	1 TAT_HV1A2	P04614 human immun
15	411	74.5	101	1 TAT_HV1OY	P08093 human immun
16	400	72.5	86	1 TAT_HV122	P12506 human immun
17	400	72.5	86	1 TAT_HV126	P04609 human immun
18	395	71.6	86	1 TAT_HV1ND	P18804 human immun
19	387.5	70.2	87	1 TAT_HV1MA	P04613 human immun
20	384	69.6	99	1 TAT_HV1EL	P04611 human immun
21	342	62.0	101	1 TAT_HV1U4	P24738 human immun
22	324	58.7	58	1 TAT_HV1B5	P04612 human immun
23	315.5	57.2	100	1 TAT_SIVCZ	P17285 chimpanzee
24	175.5	31.8	130	1 TAT_HV2G1	P18044 human immun
25	174	31.5	130	1 TAT_HV2KR	P04124 human immun
26	170.5	30.9	130	1 TAT_HV2ST	P20880 human immun
27	163.5	29.6	130	1 TAT_HV2RO	P04605 human immun
28	161.5	29.3	133	1 TAT_HV2CA	P24109 human immun
29	161	29.2	130	1 TAT_HV2D1	P17759 human immun
30	159.5	28.9	130	1 TAT_HV2M1	P05911 simian immun
31	158.5	28.7	130	1 TAT_HV2BE	P18098 human immun
32	158.5	28.7	130	1 TAT_HV2N2	P05909 human immun
33	153.5	27.8	130	1 TAT_HV2SB	P12453 human immun

## ALIGNMENTS

## RESULT 1

ID	TAT_HV1B1	STANDARD;	PRT;	86 AA.
AC	P04606;			
DT	13-AUG-1987 (Rel. 05, Created)			
DT	13-AUG-1987 (Rel. 05, Last sequence update)			
DT	16-OCT-2001 (Rel. 40, Last annotation update)			
DE	TAT protein (Transactivating regulatory protein).			
GN	TAT			
OS	Human immunodeficiency virus type 1 (BH10 isolate) (HIV-1), and			
OS	Human immunodeficiency virus type 1 (HXB3 isolate) (HIV-1).			
OC	Viruses; Retroviral viruses; Retroviridae; Lentivirus.			
OX	NCBI_TaxID=11678, 11707;			
RN	[1]			
RP	SEQUENCE FROM N.A.			
RC	STRAIN-ISOLATE BH10;			
RX	MEDLINE=8511123; PubMed=2578615;			
RA	Ratner L., Haseltine W., Patarca R., Livak K.J., Starcich B.R.,			
RA	Josephs S.F., Doran E.R., Rafalski J.A., Whitehorn E.A.,			
RA	Baumeister K., Ivanoff L., Petteway S.R. Jr., Pearson M.L.,			
RA	Lautenberger J.A., Papas T.S., Grayeb J., Chang N.T., Gallo R.C.,			
RA	Wong-Staal F.;			
RT	"Complete nucleotide sequence of the AIDS virus, HTLV-III.";			
RL	Nature 313:277-284(1985).			
RN	[2]			
RP	SEQUENCE FROM N.A.			
RC	STRAIN-ISOLATE HXB3;			
RX	MEDLINE=85228248; PubMed=2988795;			
RA	Crowl R., Ganguly K., Gordon M., Conroy R., Schaber M., Kramer R.,			
RA	Shaw G.M., Wong-Staal F., Reddy E.P.;			
RT	"HTLV-III env gene products synthesized in E. coli are recognized by			
RL	antibodies present in the sera of AIDS patients.";			
CC	Cell 41:979-986(1985).			
CC	-1- FUNCTION: TRANSCRIPTIONAL REGULATOR THAT ACTS BY BINDING TO THE			
CC	TRANS-ACTIVATING RESPONSIVE SEQUENCE (TAR) RNA ELEMENT AND			
CC	ACTIVATES TRANSCRIPTION INITIATION AND/OR ELONGATION FROM THE LTR			
CC	PROMOTER.			
CC	-1- SUBUNIT: BINDS CYCLIN T1 (BY SIMILARITY).			
CC	-1- SUBCELLULAR LOCATION: NUCLEAR; NUCLEOLAR.			
CC	-1- MISCELLANEOUS: THE TWO ISOLATES TAT SEQUENCE ARE IDENTICAL.			
CC	This SWISS-prot entry is copyright. It is produced through a collaboration			
CC	between the Swiss Institute of Bioinformatics and the EMBL Outstation -			
CC	the European Bioinformatics Institute. There are no restrictions on its			
CC	use by non-profit institutions as long as its content is in no way			
CC	modified and this statement is not removed. Usage by and for commercial			
CC	entities requires a license agreement (See <a href="http://www.isb-sib.ch/announce/">http://www.isb-sib.ch/announce/</a>			
CC	or send an email to <a href="mailto:license@isb-sib.ch">license@isb-sib.ch</a> ).			
CC	-----			
DR	EMBL; M14100; AAA44676.1; -			
DR	EMBL; M15654; AAA44199.1; -			
DR	HIV; M15654; TAT\$BHI02.			
DR	HIV; M14100; TAT\$HXB3.			
DR	InterPro; IPR001831; HIV_Tat.			
DR	Pfam; PF00539; Tat; 1.			
DR	PRINTS; PR00055; HIVTATDOMAIN.			

P11263 simian immu  
P15835 human immu  
Q02838 simian immu  
P05910 simian immu  
P19507 simian immu  
P05913 simian immu  
P27982 simian immu  
P12513 simian immu  
P27975 simian immu  
P22384 simian immu  
P36340 simian immu  
P12507 human immu

34 149 27.0 106 1 TAT\_SIVML  
35 148.5 26.9 96 1 TAT\_HV2D2  
36 147.5 26.7 94 1 TAT\_SIVAI  
37 146.5 26.5 131 1 TAT\_SIVMK  
38 143 25.9 128 1 TAT\_SIVSP  
39 142 25.7 100 1 TAT\_SIVAT  
40 139.5 25.3 119 1 TAT\_SIVAG  
41 136.5 24.7 117 1 TAT\_SIVSA  
42 135 24.5 119 1 TAT\_SIVAI  
43 127.5 23.1 115 1 TAT\_SIVGB  
44 122 22.1 116 1 TAT\_SIVAM  
45 116 21.0 44 1 TAT\_HV1BN

KW Transcription regulation; Activator; RNA-binding; Nuclear protein;  
 KW AIDS.  
 SQ SEQUENCE 86 AA; 9784 MW; 4DD609415FAF9015 CRC64;

Query Match 86.6%; Score 478; DB 1; Length 86;  
 Best Local Similarity 96.5%; Pred. No. 1.6e-39;  
 Matches 83; Conservative 2; Mismatches 1; Indels 0; Gaps 0;

Qy 1 MEPVDPRLEPKHPSQPKTACTNCYCKKCCFHCQVCFITAAIGISYGRKKRRRRPPQ 60  
 |||||  
 Db 1 MEPVDPRLEPKHPSQPKTACTNCYCKKCCFHCQVCFITAAIGISYGRKKRRRRPPQ 60  
 |||||

Qy 61 GSQTHQVSLSKQPTSQSGKEPTGPKE 86  
 |||||  
 Db 61 GSQTHQVSLSKQPTSQSGKEPTGPKE 86  
 |||||

RESULT 2  
 TAT\_HV1PV  
 ID TAT\_HV1PV STANDARD; PRT; 86 AA.  
 AC P04607;  
 DT 13-AUG-1987 (Rel. 05, Created)  
 DT 13-AUG-1987 (Rel. 05, Last sequence update)  
 DT 16-OCT-2001 (Rel. 40, Last annotation update)  
 DE TAT protein (Transactivating regulatory protein).  
 GN TAT.  
 OS Human immunodeficiency virus type 1 (PV22 isolate) (HIV-1).  
 OC Viruses; Retroviridae; Retroviridae; Lentivirus.  
 OX NCBI\_TaxID=11700;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RX MEDLINE=8511157; PubMed=2982104;  
 RA Muesing M.A., Smith D.H., Cabradilla C.D., Benton C.V., Lasky L.A.,  
 RA Capon D.J.;  
 RT "Nucleic acid structure and expression of the human  
 RT AIDS/lymphadenopathy retrovirus.";  
 RL Nature 313:450-458(1985).  
 CC -!- FUNCTION: TRANSCRIPTIONAL REGULATOR THAT ACTS BY BINDING TO THE  
 CC TRANS-ACTIVATING RESPONSIVE SEQUENCE (TAR) RNA ELEMENT AND  
 CC ACTIVATES TRANSCRIPTION INITIATION AND/OR ELONGATION FROM THE LTR  
 CC PROMOTER.  
 CC -!- SUBUNIT: BINDS CYCLIN T1 (BY SIMILARITY).  
 CC -!- SUBCELLULAR LOCATION: NUCLEAR; NUCLEOLAR.  
 CC  
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
 CC the European Bioinformatics Institute. There are no restrictions on its  
 CC use by non-profit institutions as long as its content is in no way  
 CC modified and this statement is not removed. Usage by and for commercial  
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announcement/>  
 CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
 CC  
 CC EMBL; K02083; AAB59870.1;  
 CC EMBL; X01762; -; NOT ANNOTATED\_CDS.  
 CC HIV; K02083; TATSPV22.  
 CC InterPro; IPR001831; HIV\_Tat.  
 CC Pfam; PF00539; Tat; 1.  
 CC PRINTS; PR00055; HIVTATDOMAIN.  
 CC Transcription regulation; Activator; RNA-binding; Nuclear protein;  
 KW AIDS.  
 SQ SEQUENCE 86 AA; 9794 MW; 4DD5C6415FAF9015 CRC64;

Query Match 85.7%; Score 473; DB 1; Length 86;  
 Best Local Similarity 95.3%; Pred. No. 4.9e-39;  
 Matches 82; Conservative 2; Mismatches 2; Indels 0; Gaps 0;

Qy 1 MEPVDPRLEPKHPSQPKTACTNCYCKKCCFHCQVCFITAAIGISYGRKKRRRRPPQ 60  
 |||||  
 Db 1 MEPVDPRLEPKHPSQPKTACTNCYCKKCCFHCQVCFITAAIGISYGRKKRRRRPPQ 60  
 |||||

Qy 61 GSQTHQVSLSKQPTSQSGKEPTGPKE 86  
 |||||  
 Db 61 GSQTHQVSLSKQPTSQSGKEPTGPKE 86  
 |||||

RESULT 3  
 TAT\_HV112  
 ID TAT\_HV112 STANDARD; PRT; 86 AA.  
 AC P04326;  
 DT 20-MAR-1987 (Rel. 04, Created)  
 DT 13-AUG-1987 (Rel. 05, Last sequence update)  
 DT 16-OCT-2001 (Rel. 40, Last annotation update)  
 DE TAT protein (Transactivating regulatory protein).  
 GN TAT.  
 OS Human immunodeficiency virus type 1 (clone 12) (HIV-1).  
 OC Viruses; Retroviridae; Retroviridae; Lentivirus.  
 OX NCBI\_TaxID=11679;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RX MEDLINE=86177573; PubMed=3008154;  
 RA Arya S.K., Gallo R.C.;  
 RT "Three novel genes of human T-lymphotropic virus type III: immune  
 RT reactivity of their products with sera from acquired immune  
 RT deficiency syndrome patients.";  
 RL Proc. Natl. Acad. Sci. U.S.A. 83:2209-2213(1986).  
 CC -!- FUNCTION: TRANSCRIPTIONAL REGULATOR THAT ACTS BY BINDING TO THE  
 CC TRANS-ACTIVATING RESPONSIVE SEQUENCE (TAR) RNA ELEMENT AND  
 CC ACTIVATES TRANSCRIPTION INITIATION AND/OR ELONGATION FROM THE LTR  
 CC PROMOTER.  
 CC -!- SUBUNIT: BINDS CYCLIN T1 (BY SIMILARITY).  
 CC -!- SUBCELLULAR LOCATION: NUCLEAR; NUCLEOLAR.  
 CC  
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
 CC the European Bioinformatics Institute. There are no restrictions on its  
 CC use by non-profit institutions as long as its content is in no way  
 CC modified and this statement is not removed. Usage by and for commercial  
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announcement/>  
 CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
 CC  
 CC EMBL; M11840; AAA44999.1;  
 CC PIR; A04017; TNLJ12.  
 CC HIV; M11840; TATSPCV12.  
 CC InterPro; IPR001831; HIV\_Tat.  
 CC Pfam; PF00539; Tat; 1.  
 CC PRINTS; PR00055; HIVTATDOMAIN.  
 CC Transcription regulation; Activator; RNA-binding; Nuclear protein;  
 KW AIDS.  
 SQ SEQUENCE 86 AA; 9758 MW; 4DD609414FBE9115 CRC64;

Query Match 85.1%; Score 470; DB 1; Length 86;  
 Best Local Similarity 95.3%; Pred. No. 9.5e-39;  
 Matches 82; Conservative 2; Mismatches 2; Indels 0; Gaps 0;

Qy 1 MEPVDPRLEPKHPSQPKTACTNCYCKKCCFHCQVCFITAAIGISYGRKKRRRRPPQ 60  
 |||||  
 Db 1 MEPVDPRLEPKHPSQPKTACTNCYCKKCCFHCQVCFITAAIGISYGRKKRRRRPPQ 60  
 |||||

Qy 61 GSQTHQVSLSKQPTSQSGKEPTGPKE 86  
 |||||  
 Db 61 GSQTHQVSLSKQPTSQSGKEPTGPKE 86  
 |||||

RESULT 4  
 TAT\_HV1BR  
 ID TAT\_HV1BR STANDARD; PRT; 86 AA.  
 AC P04610;  
 DT 13-AUG-1987 (Rel. 05, Created)  
 DT 13-AUG-1987 (Rel. 05, Last sequence update)  
 DT 16-OCT-2001 (Rel. 40, Last annotation update)  
 DE TAT protein (Transactivating regulatory protein).  
 GN TAT.  
 OS Human immunodeficiency virus type 1 (BRU isolate) (HIV-1).

OC Viruses; Retroviral viruses; Retroviridae; Lentivirus.  
OX NCBI\_TaxID=11686;  
RN [1]  
RP SEQUENCE FROM N.A. (CLONE PNL4-3).  
RX MEDLINE=85099333; PubMed=2981635;  
RA Wain-Hobson S., Sonigo P., Danos O., Cole S., Alizon M.;  
RT "Nucleotide sequence of the AIDS virus, LAV.";  
RL Cell 40:9-17(1985).  
RN [2]  
RN SEQUENCE FROM N.A. (CLONE PNL4-3).  
RP Buckler C.E., Buckler-White A.J., Willey R.L., McCoy J.;  
RA Submitted (JUN-1988) to the EMBL/GenBank/DDBJ databases.  
RL  
CC -!- FUNCTION: TRANSCRIPTIONAL REGULATOR THAT ACTS BY BINDING TO THE  
CC TRANS-ACTIVATING RESPONSE SEQUENCE (TAR) RNA ELEMENT AND  
CC ACTIVATES TRANSCRIPTION INITIATION AND/OR ELONGATION FROM THE LTR  
CC PROMOTER.  
CC  
CC -!- SUBUNIT: Binds cyclin T1 (BY SIMILARITY).  
CC  
CC -!- SUBCELLULAR LOCATION: Nuclear; nucleolar.  
CC -----  
CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
CC between the Swiss Institute of Bioinformatics and the EMBL Outstation -  
CC the European Bioinformatics Institute. There are no restrictions on its  
CC use by non-profit institutions as long as its content is in no way  
CC modified and this statement is not removed. Usage by and for commercial  
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/  
CC or send an email to license@isb-sib.ch).  
CC -----  
CC EMBL; K02013; AA959745.1; -  
DR EMBL; M19921; AAA44985.1; -  
DR HIV; K02013; TATSRU.  
DR HIV; M19921; TATSRL43.  
DR InterPro; IPR001831; HIV\_Tat.  
DR Pfam; PF00539; Tat; 1.  
DR PRINTS; PR00055; HVTATDOMAIN.  
DR Transcription regulation; Activator; RNA-binding; Nuclear protein;  
KW AIDS.  
KW VARIANT 24 24 T->N (IN CLONE PNL4-3).  
FT VARIANT 29 39 T->M (IN CLONE PNL4-3).  
FT VARIANT 58 61 PPQG->AHQN (IN CLONE PNL4-3).  
FT VARIANT 67 67 V->A (IN CLONE PNL4-3).  
FT VARIANT 77 77 P->S (IN CLONE PNL4-3).  
SQ SEQUENCE 86 AA; 9769 MW; 9B1B4A915FAF8A14 CRC64;

Query Match 83.7%; Score 462; DB 1; Length 86;  
Best Local Similarity 93.0%; Pred. No. 5.5e-38;  
Matches 80; Conservative 2; Mismatches 4; Indels 0; Gaps 0;  
  
QY 1 MEPVDPRLPEPKHGPSQPACTCYCKKCFHCQCVCFFITAAAGISGGRKKRRRPPQ 60  
| | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |  
Db 1 MEPVDPRLPEPKHGPSQPACTCYCKKCFHCQCVCFFITAAAGISGGRKKRRRPPQ 60  
  
QY 61 GSOTHVSLSKSQTSSKGPTGPKE 86  
| | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |  
Db 61 GSOTHVSLSKSQTSSKGPTGPKE 86  
  
RESULT 5  
TAT\_HVIH STANDARD; PRT; 102 AA.  
AC P05508;  
DT 01-NOV-1988 (Rel. 09, Created)  
DT 01-NOV-1988 (Rel. 09, Last sequence update)  
DT 16-OCT-2001 (Rel. 40, Last annotation update)  
DE TAT protein (Transactivating regulatory protein).  
GN TAT.  
OS Human immunodeficiency virus type 1 (RF/HAT isolate) (HTV-1).  
OC Viruses; Retroviral viruses; Retroviridae; Lentivirus.  
OX NCBI\_TaxID=11701;  
RN [1]  
RP SEQUENCE FROM N.A.  
RX MEDLINE=86218077; PubMed=2423250;  
RA Starcich B.R., Hann B.H., Shaw G.M., McNeely P.D., Modrow S.

[illegible]

```
CC -!- SUBUNIT: BINDS CYCLIN T1 (BY SIMILARITY).
CC -!- SUBCELLULAR LOCATION: NUCLEAR; NUCLEOLAR.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL; K03455; AAB50256.1; -.
DR HIV; AF033819; AAC82591.1; -.
DR HIV; K03455; TAT$HXB2.
DR InterPro; IPR001831; HIV_Tat.
DR Pfam; PF00539; Tat; 1.
DR PRINTS; PR00055; HIVTATDOMAIN.
DR Transcription regulation; Activator; RNA-binding; Nuclear protein;
KW AIDS.
KW SEQUENCE 86 AA; 9837 MW; 4DC56D979769115 CRC64;
-----
Query Match 80.8%; Score 446; DB 1; Length 86;
Best Local Similarity 90.7%; Pred. No. 1.9e-36;
Matches 78; Conservative 2; Mismatches 6; Indels 0; Gaps 0;
Qy 1 MEPVDPRLEPKHGPSQPKTACTNCKYCKKCCFHCQVCFITAAALGISYGRKKRRRRPPQ 60
Db 1 MEPVDPRLEPKHGPSQPKTACTNCKYCKKCCFHCQVCFITKALGISYGRKKRRRAHQ 60
Qy 61 GSQTHQVSLSKQPTSQSKGEPTGPK 86
Db 61 NSQTHQASLSKQPTSQPRGDPGPK 86
-----
RESULT 7
TAT_HV1JR
ID TAT_HV1JR STANDARD; PRT; 101 AA.
AC P20879;
DT 01-FEB-1991 (Rel. 17, Created)
DT 01-FEB-1991 (Rel. 17, Last sequence update)
DT 16-OCT-2001 (Rel. 40, Last annotation update)
DE TAT protein (Transactivating regulatory protein).
GN TAT.
OS Human immunodeficiency virus type 1 (JRCSF isolate) (HIV-1).
OC Viruses; Retroviridae; Lentivirus.
OX NCBI_TaxID=11686;
RN [1]
RP SEQUENCE FROM N.A.
RA Koyanagi S., Chen I.S.Y.;
RL Submitted (DEC-1988) to the HIV data bank.
CC -!- FUNCTION: TRANSCRIPTIONAL REGULATOR THAT ACTS BY BINDING TO THE
CC TRANS-ACTIVATING RESPONSIVE SEQUENCE (TAR) RNA ELEMENT AND
CC ACTIVATES TRANSCRIPTION INITIATION AND/OR ELONGATION FROM THE LTR
CC PROMOTER.
CC -!- SUBUNIT: BINDS CYCLIN T1 (BY SIMILARITY).
CC -!- SUBCELLULAR LOCATION: NUCLEAR; NUCLEOLAR.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL; M38429; -. NOT_ANNOTATED_CDS.
DR HIV; M38429; TAT$JRCSF.
DR InterPro; IPR001831; HIV_Tat.
DR Pfam; PF00539; Tat; 1.
DR PRINTS; PR00055; HIVTATDOMAIN.
DR Transcription regulation; Activator; RNA-binding; Nuclear protein;
KW AIDS.
```

```
SQ SEQUENCE 101 AA; 11508 MW; 3E6CBD8FF7F8D4FA CRC64;
-----
Query Match 79.3%; Score 438; DB 1; Length 101;
Best Local Similarity 86.2%; Pred. No. 1.3e-35;
Matches 75; Conservative 4; Mismatches 8; Indels 0; Gaps 0;
Qy 1 MEPVDPRLEPKHGPSQPKTACTNCKYCKKCCFHCQVCFITAAALGISYGRKKRRRRPPQ 60
Db 1 MEPVDPRLEPKHGPSQPKTACTNCKYCKKCCFHCQVCFITKGLGISYGRKKRRRRPPQ 60
Qy 61 GSQTHQVSLSKQPTSQSKGEPTGPK 87
Db 61 DSQTHQVSLPKQPSQOGRDPTGPKES 87
-----
RESULT 8
TAT_HV1S1
ID TAT_HV1S1 STANDARD; PRT; 101 AA.
AC P19553;
DT 01-FEB-1991 (Rel. 17, Created)
DT 01-FEB-1991 (Rel. 17, Last sequence update)
DT 16-OCT-2001 (Rel. 40, Last annotation update)
DE TAT protein (Transactivating regulatory protein).
GN TAT.
OS Human immunodeficiency virus type 1 (SF162 isolate) (HIV-1).
OC Viruses; Retroviridae; Lentivirus.
OX NCBI_TaxID=11691;
RN [1]
RP SEQUENCE FROM N.A.
RA MEDLINE=90347835; PubMed=2384920;
RA Cheng-Mayer C., Quiroga M., Tung J.W., Dina D., Levy J.;
RT "Viral determinants of human immunodeficiency virus type 1 T-cell or
RT macrophage tropism, cytopathogenicity, and CD4 antigen modulation.";
RL J. Virol. 64:4390-4398(1990).
CC -!- FUNCTION: TRANSCRIPTIONAL REGULATOR THAT ACTS BY BINDING TO THE
CC TRANS-ACTIVATING RESPONSIVE SEQUENCE (TAR) RNA ELEMENT AND
CC ACTIVATES TRANSCRIPTION INITIATION AND/OR ELONGATION FROM THE LTR
CC PROMOTER.
CC -!- SUBUNIT: BINDS CYCLIN T1 (BY SIMILARITY).
CC -!- SUBCELLULAR LOCATION: NUCLEAR; NUCLEOLAR.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL; M65024; AAA45069.1; -.
DR HIV; M38428; TAT$SF162.
DR InterPro; IPR001831; HIV_Tat.
DR Pfam; PF00539; Tat; 1.
DR PRINTS; PR00055; HIVTATDOMAIN.
DR Transcription regulation; Activator; RNA-binding; Nuclear protein;
KW AIDS.
KW SEQUENCE 101 AA; 11497 MW; 8CE2C48C2C63293C CRC64;
-----
Query Match 77.9%; Score 430; DB 1; Length 101;
Best Local Similarity 85.1%; Pred. No. 7.4e-35;
Matches 74; Conservative 4; Mismatches 9; Indels 0; Gaps 0;
Qy 1 MEPVDPRLEPKHGPSQPKTACTNCKYCKKCCFHCQVCFITAAALGISYGRKKRRRRPPQ 60
Db 1 MEPVDPRLEPKHGPSQPKTACTNCKYCKKCCFHCQVCFITKGLGISYGRKKRRRRAPP 60
Qy 61 GSQTHQVSLSKQPTSQSKGEPTGPK 87
Db 61 DSEVHQVSLPKQPSQOGRDPTGPKES 87
```

```
RESULT 9
TAT_HV1SC
ID TAT_HV1SC STANDARD; PRT; 101 AA.
AC P05906;
DT 01-NOV-1988 (Rel. 09, Created)
DT 01-NOV-1988 (Rel. 09, Last sequence update)
DT 16-OCT-2001 (Rel. 40, Last annotation update)
DE TAT protein (Transactivating regulatory protein).
GN TAT.
OS Human immunodeficiency virus type 1 (SC isolate) (HIV-1).
OC Viruses; Retrovirdae; Lentivirus.
OX NCBI_TaxID=11702;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=88219542; PubMed=3369091;
RA Gurgu C., Guo H.-G., Franchini G., Aldovini A., Collalti E.,
RA Farrell K., Wong-Staal F., Gallo R.C., Reitz M.S. Jr.;
RT "Envelope sequences of two new United States HIV-1 isolates.";
RL Virology 164:531-536(1988).
CC -1- FUNCTION: TRANSCRIPTIONAL REGULATOR THAT ACTS BY BINDING TO THE
CC TRANS-ACTIVATING RESPONSIVE SEQUENCE (TAR) RNA ELEMENT AND
CC ACTIVATES TRANSCRIPTION INITIATION AND/OR ELONGATION FROM THE LTR
CC PROMOTER.
CC -1- SUBUNIT: BINDS CYCLIN T1 (BY SIMILARITY).
CC -1- SUBCELLULAR LOCATION: NUCLEAR; NUCLEOLAR.
CC [1]
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL Outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
CC EMBL; M17450; AAA45060.1; -
CC HIV; M17450; TATSSC.
CC InterPro: IPR001831; HIV_Tat.
CC Pfam; PF00539; Tat; 1.
CC PRINTS; PR00055; HIVTATDOMAIN.
CC Transcription regulation; Activator; RNA-binding; Nuclear protein;
CC AIDS.
CC SEQUENCE 101 AA; 11426 MW; C91090B938D15788 CRC64;
-----
Query Match 77.7%; Score 429; DB 1; Length 101;
Best Local Similarity 83.9%; Pred. No. 9.2e-35;
Matches 73; Conservative 6; Mismatches 8; Indels 0; Gaps 0;
Qy 1 MEPVDPRLPKHPGSPQKPTACTNCYCKKCCFHCQVCFITAAALGYSYGRKKRRRPPQ 60
Db 1 MDPVDPRLPKHPGSPQKAACTSCYCKKCCFHCQVCFITKGLGYSYGRKKRRRAPQ 60
Qy 61 GSQTHQVSLSKOPTSQSKGPTGPKET 87
Db 61 DSQTHQVSLPKOPASQARGDPTGPRES 87
-----
RESULT 10
TAT_HV1C4
ID TAT_HV1C4 STANDARD; PRT; 101 AA.
AC P05907;
DT 01-NOV-1988 (Rel. 09, Created)
DT 01-NOV-1988 (Rel. 09, Last sequence update)
DT 16-OCT-2001 (Rel. 40, Last annotation update)
DE TAT protein (Transactivating regulatory protein).
GN TAT.
OS Human immunodeficiency virus type 1 (CDC-451 isolate) (HIV-1).
OC Viruses; Retrovirdae; Lentivirus.
OX NCBI_TaxID=11687;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=93021387; PubMed=1404605;
RA Li Y., Hui H., Burgess C.J., Price R.W., Sharp P.M., Hahn B.H.,
RA Shaw G.M.;
RT "Complete nucleotide sequence, genome organization, and biological
RT properties of human immunodeficiency virus type 1 in vivo: evidence
RT for limited defectiveness and complementation.";
RL J. Virol. 66:6587-6600(1992).
CC -1- FUNCTION: TRANSCRIPTIONAL REGULATOR THAT ACTS BY BINDING TO THE
CC TRANS-ACTIVATING RESPONSIVE SEQUENCE (TAR) RNA ELEMENT AND
CC ACTIVATES TRANSCRIPTION INITIATION AND/OR ELONGATION FROM THE LTR
CC PROMOTER.
CC -1- SUBUNIT: BINDS CYCLIN T1 (BY SIMILARITY).
CC -1- SUBCELLULAR LOCATION: NUCLEAR; NUCLEOLAR.
```

```

RX MEDLINE=87041461; PubMed=3490666;
RA Desai S.M., Kalyanaraman V.S., Casey J.M., Srinivasan A.,
RA Andersen P.R., Devare S.G.;
RT "Molecular cloning and primary nucleotide sequence analysis of a
RT distinct human immunodeficiency virus isolate reveal significant
RT divergence in its genomic sequences.";
RL Proc. Natl. Acad. Sci. U.S.A. 83:8380-8384(1986).
CC -1- FUNCTION: TRANSCRIPTIONAL REGULATOR THAT ACTS BY BINDING TO THE
CC TRANS-ACTIVATING RESPONSIVE SEQUENCE (TAR) RNA ELEMENT AND
CC ACTIVATES TRANSCRIPTION INITIATION AND/OR ELONGATION FROM THE LTR
CC PROMOTER.
CC -1- SUBUNIT: BINDS CYCLIN T1 (BY SIMILARITY).
CC -1- SUBCELLULAR LOCATION: NUCLEAR; NUCLEOLAR.
CC [1]
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL Outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
CC EMBL; M13137; AAA44309.1; -
CC HIV; M13137; TATSCDC45.
CC InterPro: IPR001831; HIV_Tat.
CC Pfam; PF00539; Tat; 1.
CC PRINTS; PR00055; HIVTATDOMAIN.
CC Transcription regulation; Activator; RNA-binding; Nuclear protein;
CC AIDS.
CC SEQUENCE 101 AA; 11594 MW; 817D915F3FB1C7FA CRC64;
-----
Query Match 77.4%; Score 427; DB 1; Length 101;
Best Local Similarity 86.0%; Pred. No. 1.4e-34;
Matches 74; Conservative 4; Mismatches 8; Indels 0; Gaps 0;
Qy 1 MEPVDPRLPKHPGSPQKPTACTNCYCKKCCFHCQVCFITAAALGYSYGRKKRRRPPQ 60
Db 1 MEPVDPRLPKHPGSPQKPTACTNCYCKKCCFHCQVCFITKALGYSYGRKKRRRAHQ 60
Qy 61 GSQTHQVSLSKOPTSQSKGPTGPK 86
Db 61 DSQNHQASLSKQSPSQTRGDPTGPKE 86
-----
RESULT 11
TAT_HV1Y2
ID TAT_HV1Y2 STANDARD; PRT; 101 AA.
AC P35965;
DT 01-JUN-1994 (Rel. 29, Created)
DT 01-JUN-1994 (Rel. 29, Last sequence update)
DT 16-OCT-2001 (Rel. 40, Last annotation update)
DE TAT protein (Transactivating regulatory protein).
GN TAT.
OS Human immunodeficiency virus type 1 (YU-2 isolate) (HIV-1).
OC Viruses; Retrovirdae; Lentivirus.
OX NCBI_TaxID=36377;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=93021387; PubMed=1404605;
RA Li Y., Hui H., Burgess C.J., Price R.W., Sharp P.M., Hahn B.H.,
RA Shaw G.M.;
RT "Complete nucleotide sequence, genome organization, and biological
RT properties of human immunodeficiency virus type 1 in vivo: evidence
RT for limited defectiveness and complementation.";
RL J. Virol. 66:6587-6600(1992).
CC -1- FUNCTION: TRANSCRIPTIONAL REGULATOR THAT ACTS BY BINDING TO THE
CC TRANS-ACTIVATING RESPONSIVE SEQUENCE (TAR) RNA ELEMENT AND
CC ACTIVATES TRANSCRIPTION INITIATION AND/OR ELONGATION FROM THE LTR
CC PROMOTER.
CC -1- SUBUNIT: BINDS CYCLIN T1 (BY SIMILARITY).
CC -1- SUBCELLULAR LOCATION: NUCLEAR; NUCLEOLAR.
```





```
Db 61 SSQHDSLSKQPSQSRGDPGPKES 87

RESULT 14
TAT_HV1A2
ID TAT_HV1A2 STANDARD; PRT; 101 AA.
AC P04614;
DT 13-AUG-1987 (Rel. 05, Created)
DT 13-AUG-1987 (Rel. 05, Last sequence update)
DT 16-OCT-2001 (Rel. 40, Last annotation update)
DE TAT protein (transactivating regulatory protein).
GN TAT.
OS Human immunodeficiency virus type 1 (ARV2/SF2 isolate) (HIV-1).
OC Viruses; Retrovirdae; Lentivirus.
OX NCBI_TaxID=11685;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=85090453; PubMed=2578227;
RA Sanchez-Pescador R., Power M.D., Barr P.J., Steimer K.S.,
RA Stempien M.M., Brown-Shiner S.L., Gee W.W., Renard A., Randolph A.,
RA Levy J.A., Dina D., Luciw P.A.;
RT Nucleotide sequence and expression of an AIDS-associated retrovirus
RT (ARV-2).;
RL Science 227:484-492(1985).
CC -1- FUNCTION: TRANSCRIPTIONAL REGULATOR THAT ACTS BY BINDING TO THE
CC TRANS-ACTIVATING RESPONSIVE SEQUENCE (TAR) RNA ELEMENT AND
CC ACTIVATES TRANSCRIPTION INITIATION AND/OR ELONGATION FROM THE LTR
CC PROMOTER.
CC -1- SUBUNIT: BINDS CYCLIN T1 (BY SIMILARITY).
CC -1- SUBCELLULAR LOCATION: NUCLEAR; NUCLEOLAR.
CC -1- MISCELLANEOUS: THE OVI ISOLATE WAS TAKEN FROM THE BLOOD OF A
CC HEALTHY GABONESE INDIVIDUAL.
CC -1- MISCELLANEOUS: THE SINGLE C -> S SUBSTITUTION AT RESIDUE 22 OF THE
CC OVI TAT PROTEIN RENDERS IT INACTIVE, BUT MAY NOT ACCOUNT FOR THE
CC AVIRULENCE OF THE VIRUS.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
CC EMBL; M26727; AA83395.1; -
CC HIV; M26727; TATSOYI.
CC InterPro; IPR001831; HIV_Tat.
CC Pfam; PF00539; Tat; 1.
CC PRINTS; PR00055; HIVTATDOMAIN.
CC Transcription regulation; Activator; RNA-binding; Nuclear protein;
CC AIDS.
CC SEQUENCE 101 AA; 11557 MW; 1762370A3BD641FD CRC64;

Query Match 74.5%; Score 411; DB 1; Length 101;
Best Local Similarity 80.5%; Pred. No. 4.9e-33;
Matches 70; Conservative 4; Mismatches 13; Indels 0; Gaps 0;

Qy 1 MEPVDPRLEPKHFGSQPKTACTNCKKCCFHCQVCFITAAALGISYGRKKRRRRPPQ 60
Db 1 MEPVDNPLEPKHFGSQPKTACNCKKCCFHCQVCFITAAALGISYGRKKRRRRPPQ 60
Qy 61 GSQTHQVSLSKQPTSQSGKEPTGPKET 87
Db 61 DSQTHQASLSKQPSQSRGDPGPKES 87

RESULT 15
TAT_HV1OY
ID TAT_HV1OY STANDARD; PRT; 101 AA.
AC P20893;
DT 01-FEB-1991 (Rel. 17, Created)
DT 01-FEB-1991 (Rel. 17, Last sequence update)
DT 16-OCT-2001 (Rel. 40, Last annotation update)
DE TAT protein (transactivating regulatory protein).
GN TAT.
OS Human immunodeficiency virus type 1 (OVI isolate) (HIV-1).
OC Viruses; Retrovirdae; Lentivirus.
OX NCBI_TaxID=11683;
RN [1]
RP SEQUENCE FROM N.A.
RA Theodore T., Buckler-White A.;
RL Submitted (NOV-1988) to the HIV data bank.
CC -1- FUNCTION: TRANSCRIPTIONAL REGULATOR THAT ACTS BY BINDING TO THE
CC TRANS-ACTIVATING RESPONSIVE SEQUENCE (TAR) RNA ELEMENT AND
CC ACTIVATES TRANSCRIPTION INITIATION AND/OR ELONGATION FROM THE LTR
CC PROMOTER.
CC -1- SUBUNIT: BINDS CYCLIN T1 (BY SIMILARITY).
CC -1- SUBCELLULAR LOCATION: NUCLEAR; NUCLEOLAR.
```

```
OX NCBI_TaxID=11699;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=90148544; PubMed=2559749;
RA Huet T., Dazza M.C., Brun-Vezinet F., Roelants G.E., Wain-Hobson S.;
RT "A highly defective HIV-1 strain isolated from a healthy Gabonese
RT individual presenting an atypical western blot.";
RL AIDS 3:707-715(1989).
CC -1- FUNCTION: TRANSCRIPTIONAL REGULATOR THAT ACTS BY BINDING TO THE
CC TRANS-ACTIVATING RESPONSIVE SEQUENCE (TAR) RNA ELEMENT AND
CC ACTIVATES TRANSCRIPTION INITIATION AND/OR ELONGATION FROM THE LTR
CC PROMOTER.
CC -1- SUBUNIT: BINDS CYCLIN T1 (BY SIMILARITY).
CC -1- SUBCELLULAR LOCATION: NUCLEAR; NUCLEOLAR.
CC -1- MISCELLANEOUS: THE OVI ISOLATE WAS TAKEN FROM THE BLOOD OF A
CC HEALTHY GABONESE INDIVIDUAL.
CC -1- MISCELLANEOUS: THE SINGLE C -> S SUBSTITUTION AT RESIDUE 22 OF THE
CC OVI TAT PROTEIN RENDERS IT INACTIVE, BUT MAY NOT ACCOUNT FOR THE
CC AVIRULENCE OF THE VIRUS.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
CC EMBL; M26727; AA83395.1; -
CC HIV; M26727; TATSOYI.
CC InterPro; IPR001831; HIV_Tat.
CC Pfam; PF00539; Tat; 1.
CC PRINTS; PR00055; HIVTATDOMAIN.
CC Transcription regulation; Activator; RNA-binding; Nuclear protein;
CC AIDS.
CC SEQUENCE 101 AA; 11561 MW; 83ACAC36CC4C87AD CRC64;

Query Match 74.5%; Score 411; DB 1; Length 101;
Best Local Similarity 81.6%; Pred. No. 4.9e-33;
Matches 71; Conservative 5; Mismatches 11; Indels 0; Gaps 0;

Qy 1 MEPVDPRLEPKHFGSQPKTACTNCKKCCFHCQVCFITAAALGISYGRKKRRRRPPQ 60
Db 1 MEPVDPRLEPKHFGSQPKTASNNCKKCCFHCQVCFITAAALGISYGRKKRRRRPPQ 60
Qy 61 GSQTHQVSLSKQPTSQSGKEPTGPKET 87
Db 61 DSKTHQVSLSKQPSQSRGDPGPKES 87

RESULT 16
TAT_HV122
ID TAT_HV122 STANDARD; PRT; 86 AA.
AC P12506;
DT 01-OCT-1989 (Rel. 12, Created)
DT 01-OCT-1989 (Rel. 12, Last sequence update)
DT 16-OCT-2001 (Rel. 40, Last annotation update)
DE TAT protein (transactivating regulatory protein).
GN TAT.
OS Human immunodeficiency virus type 1 (22/CDC-234 isolate) (HIV-1).
OC Viruses; Retrovirdae; Lentivirus.
OX NCBI_TaxID=11683;
RN [1]
RP SEQUENCE FROM N.A.
RA Theodore T., Buckler-White A.;
RL Submitted (NOV-1988) to the HIV data bank.
CC -1- FUNCTION: TRANSCRIPTIONAL REGULATOR THAT ACTS BY BINDING TO THE
CC TRANS-ACTIVATING RESPONSIVE SEQUENCE (TAR) RNA ELEMENT AND
CC ACTIVATES TRANSCRIPTION INITIATION AND/OR ELONGATION FROM THE LTR
CC PROMOTER.
CC -1- SUBUNIT: BINDS CYCLIN T1 (BY SIMILARITY).
CC -1- SUBCELLULAR LOCATION: NUCLEAR; NUCLEOLAR.
```

-----  
This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation - the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See <http://www.isb-sib.ch/announcement/> or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
-----  
EMBL; M22639; AAA45363.1; -  
HIV; M22639; TAT\$2226.  
InterPro: IPR001831; HIV\_Tat.  
Pfam: PF00539; Tat; 1.  
PRINTS; PR00055; HIVTATDOMAIN  
Transcription regulation; Activator; RNA-binding; Nuclear protein;  
AIDS.  
KW AIDS.  
SQ SEQUENCE 86 AA; 9737 MW; 97561D72AFCFF19 CRC64;  
-----  
Query Match 72.5%; Score 400; DB 1; Length 86;  
Best Local Similarity 76.7%; Pred. No. 4.8e-32;  
Matches 66; Conservative 8; Mismatches 12; Indels 0; Gaps 0;  
-----  
QY 1 MEPVDPRLEPKHGPSQPKTACTNCYCKKCFHCQVCFITAAALGISYGRKKRRRPPQ 60  
I:||||:||||:||||:||||:||||:||||:||||:||||:||||:||||:||||:|  
DB 1 MDPVDNPEPNHGPSQPKTACNRCHCKKCCYHCQVCFITKGLGISYGRKKRRRPPQ 60  
I:||||:||||:||||:||||:||||:||||:||||:||||:||||:||||:|  
QY 61 GSQTHQVSLSKOPTSQSKGPTGPKE 86  
I:||||:||||:||||:||||:||||:||||:||||:||||:||||:||||:|  
DB 61 GGQTHQDPIPKOPSSQPRGDTGPKE 86  
I:||||:||||:||||:||||:||||:||||:||||:||||:||||:||||:|  
-----  
RESULT 17  
TAT\_HV126  
ID TAT\_HV126 STANDARD; PRT; 86 AA.  
AC P04609;  
DT 13-AUG-1987 (Rel. 05, Created)  
DT 13-AUG-1987 (Rel. 05, Last sequence update)  
DT 16-OCT-2001 (Rel. 40, Last annotation update)  
DE TAT protein (Transactivating regulatory protein).  
GN Tat.  
OS Human immunodeficiency virus type 1 (Zaire 6 isolate) (HIV-1).  
OC Viruses; Retroid viruses; Retroviridae; Lentivirus.  
OX NCBI\_TaxID=11708;  
RN [1]  
RP SEQUENCE FROM N.A.  
RX MEDLINE=87248097; PubMed=3036660;  
RA Srinivasan A., Anand R., York D., Ranganathan P., Feorino P.,  
RA Schochetman G., Curran J., Kalyanaram V.S., Luciw P.A.,  
RA Sanchez-Pescador R.;  
RT "Molecular characterization of human immunodeficiency virus from  
RT Zaire: nucleotide sequence analysis identifies conserved and variable  
RT domains in the envelope gene";  
RL Gene 52:71-82(1987).  
CC -!- FUNCTION: TRANSCRIPTIONAL REGULATOR THAT ACTS BY BINDING TO THE  
CC TRANS-ACTIVATING RESPONSIVE SEQUENCE (TAR) RNA ELEMENT AND  
CC ACTIVATES TRANSCRIPTION INITIATION AND/OR ELONGATION FROM THE LTR  
CC PROMOTER.  
CC -!- SUBUNIT: BINDS CYCLIN T1 (BY SIMILARITY).  
CC -!- SUBCELLULAR LOCATION: NUCLEAR; NUCLEOLAR.  
CC -!- MISCELLANEOUS: NDK, ISOLATED FROM A ZAIRIAN PATIENT AFFECTED WITH  
CC AIDS, AND IS A HIGHLY CYTOPATHOGENIC STRAIN.  
CC -----  
CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
CC the European Bioinformatics Institute. There are no restrictions on its  
CC use by non-profit institutions as long as its content is in no way  
CC modified and this statement is not removed. Usage by and for commercial  
CC entities requires a license agreement (See <http://www.isb-sib.ch/announcement/>  
CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
CC -----  
EMBL; K03458; AAA45377.1; -  
DR PIR; C26192; TNLJZR.  
DR HIV; K03458; TAT\$26.  
DR InterPro: IPR001831; HIV\_Tat.  
-----

DR Pfam: PF00539; Tat; 1.  
DR PRINTS; PR00055; HIVTATDOMAIN.  
KW Transcription regulation; Activator; RNA-binding; Nuclear protein;  
KW AIDS.  
SQ SEQUENCE 86 AA; 9736 MW; E1E0AD741FCFF5AC CRC64;  
-----  
Query Match 72.5%; Score 400; DB 1; Length 86;  
Best Local Similarity 77.9%; Pred. No. 4.8e-32;  
Matches 67; Conservative 6; Mismatches 13; Indels 0; Gaps 0;  
-----  
QY 1 MEPVDPRLEPKHGPSQPKTACTNCYCKKCFHCQVCFITAAALGISYGRKKRRRPPQ 60  
I:||||:||||:||||:||||:||||:||||:||||:||||:||||:||||:|  
DB 1 MDPVDNPEPNHGPSQPKTACNRCHCKKCCYHCQVCFITKGLGISYGRKKRRRPPQ 60  
I:||||:||||:||||:||||:||||:||||:||||:||||:||||:||||:|  
QY 61 GSQTHQVSLSKOPTSQSKGPTGPKE 86  
I:||||:||||:||||:||||:||||:||||:||||:||||:||||:||||:|  
DB 61 GGQTHQDPIPKOPSSQPRGDTGPKE 86  
I:||||:||||:||||:||||:||||:||||:||||:||||:||||:||||:|  
-----  
RESULT 18  
TAT\_HVIND  
ID TAT\_HVIND STANDARD; PRT; 86 AA.  
AC P18804;  
DT 01-NOV-1990 (Rel. 16, Created)  
DT 01-NOV-1990 (Rel. 16, Last sequence update)  
DT 16-OCT-2001 (Rel. 40, Last annotation update)  
DE TAT protein (Transactivating regulatory protein).  
GN Tat.  
OS Human immunodeficiency virus type 1 (NDK isolate) (HIV-1).  
OC Viruses; Retroid viruses; Retroviridae; Lentivirus.  
OX NCBI\_TaxID=11695;  
RN [1]  
RP SEQUENCE FROM N.A.  
RX MEDLINE=90034200; PubMed=2806917;  
RA Spire B., Sire J., Zachar V., Rey F., Barre-Sinoussi F., Galibert F.,  
RA Hampe A., Hermann J.C.;  
RT "Nucleotide sequence of HIV1-NDK: a highly cytopathic strain of the  
RT human immunodeficiency virus";  
RL Gene 81:275-284(1989).  
CC -!- FUNCTION: TRANSCRIPTIONAL REGULATOR THAT ACTS BY BINDING TO THE  
CC TRANS-ACTIVATING RESPONSIVE SEQUENCE (TAR) RNA ELEMENT AND  
CC ACTIVATES TRANSCRIPTION INITIATION AND/OR ELONGATION FROM THE LTR  
CC PROMOTER.  
CC -!- SUBUNIT: BINDS CYCLIN T1 (BY SIMILARITY).  
CC -!- SUBCELLULAR LOCATION: NUCLEAR; NUCLEOLAR.  
CC -!- MISCELLANEOUS: NDK, ISOLATED FROM A ZAIRIAN PATIENT AFFECTED WITH  
CC AIDS, AND IS A HIGHLY CYTOPATHOGENIC STRAIN.  
CC -----  
CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
CC the European Bioinformatics Institute. There are no restrictions on its  
CC use by non-profit institutions as long as its content is in no way  
CC modified and this statement is not removed. Usage by and for commercial  
CC entities requires a license agreement (See <http://www.isb-sib.ch/announcement/>  
CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
CC -----  
EMBL; M27323; AAA44866.1; -  
DR PIR; JQ0071; TNLJND.  
DR HIV; M27323; TAT\$NDK.  
DR InterPro: IPR001831; HIV\_Tat.  
DR Pfam: PF00539; Tat; 1.  
DR PRINTS; PR00055; HIVTATDOMAIN.  
KW Transcription regulation; Activator; RNA-binding; Nuclear protein;  
KW AIDS.  
SQ SEQUENCE 86 AA; 9711 MW; 7DB9E64E0AF8B0F4 CRC64;  
-----  
Query Match 71.6%; Score 395; DB 1; Length 86;  
Best Local Similarity 74.4%; Pred. No. 1.4e-31;  
Matches 64; Conservative 11; Mismatches 11; Indels 0; Gaps 0;  
-----  
QY 1 MEPVDPRLEPKHGPSQPKTACTNCYCKKCFHCQVCFITAAALGISYGRKKRRRPPQ 60

```
Db 1 MDVDPNLESHNPGSQRTACNCKCHCKCCYHCQVCFITKGLGISYGRKKRRRPPQ 60
Qy 61 GSQTHQVLSKOPTSQSGKEPTGPK 86
Db 61 GQAHQVPIPEQPSQSRGDPGPK 86

RESULT 19
TAT_HV1MA
ID TAT_HV1MA STANDARD; PRT; 87 AA.
AC P04613;
DT 13-AUG-1987 (Rel. 05, Created)
DT 13-AUG-1987 (Rel. 05, Last sequence update)
DT 16-OCT-2001 (Rel. 40, Last annotation update)
DE TAT protein (Transactivating regulatory protein).
GN TAT.
OS Human immunodeficiency virus type 1 (MAL isolate) (HIV-1).
OC Viruses; Retrovirdae; Lentivirus.
OX NCBI_TaxID=11697;
RN [1]
RP MEDLINE=86245056; PubMed=2424612;
RA Allison M., Main-Hobson S., Montagnier L., Sonigo P.;
RT "Genetic variability of the AIDS virus; nucleotide sequence analysis
of two isolates from African patients.";
RL Cell 46:63-74(1986).
CC TRANS-ACTIVATING REGULATORY THAT ACTS BY BINDING TO THE
CC TRANS-ACTIVATING RESPONSE SEQUENCE (TAR) RNA ELEMENT AND
CC ACTIVATES TRANSCRIPTION INITIATION AND/OR ELONGATION FROM THE LTR
CC PROMOTER.
CC -1- SUBUNIT: BINDS CYCLIN T1 (BY SIMILARITY).
CC -1- SUBCELLULAR LOCATION: NUCLEAR; NUCLEOLAR.
CC
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC
CC EMBL; X04415; CAA28015.1; -
CC HIV; K03456; TATSMAL.
CC InterPro: IPR001831; HIV_Tat.
CC Pfam; PF00539; Tat; 1.
CC PRINTS; PR00055; HIVTATDOMAIN.
CC Transcription regulation; Activator; RNA-binding; Nuclear protein;
CC AIDS.
CC
CC SEQUENCE 87 AA; 10036 MW; 3832412849D5BICE CRC64;
```

```
-Query Match 70.2%; Score 387.5; DB 1; Length 87;
Best Local Similarity 74.7%; Pred. No. 7.6e-31;
Matches 65; Conservative 9; Mismatches 12; Indels 1; Gaps 1;

Qy 1 MEPVDPRLEPHKPGSQRTACTNCKCHCKCCYHCQVCFITKGLGISYGRKKRRRPPQ 60
Db 1 MDVDPNLEPHNPGSQRTPCNCKCHCKCCYHCQVCFITKGLGISYGRKKRRRPPQ 60
Qy 61 GSQTHQVLSKOPTSQSGKE-PTGPK 86
Db 61 GQAHQVPIPEQPSQSRGDPGPK 87

RESULT 20
TAT_HV1EL
ID TAT_HV1EL STANDARD; PRT; 99 AA.
AC P04611;
DT 13-AUG-1987 (Rel. 05, Created)
DT 13-AUG-1987 (Rel. 05, Last sequence update)
DT 16-OCT-2001 (Rel. 40, Last annotation update)
DE TAT protein (Transactivating regulatory protein).
```

```
GN TAT.
OS Human immunodeficiency virus type 1 (ELI isolate) (HIV-1).
OC Viruses; Retrovirdae; Lentivirus.
OX NCBI_TaxID=11689;
RN [1]
RP MEDLINE=86245056; PubMed=2424612;
RA Allison M., Main-Hobson S., Montagnier L., Sonigo P.;
RT "Genetic variability of the AIDS virus; nucleotide sequence analysis
of two isolates from African patients.";
RL Cell 46:63-74(1986).
CC TRANS-ACTIVATING REGULATORY THAT ACTS BY BINDING TO THE
CC TRANS-ACTIVATING RESPONSE SEQUENCE (TAR) RNA ELEMENT AND
CC ACTIVATES TRANSCRIPTION INITIATION AND/OR ELONGATION FROM THE LTR
CC PROMOTER.
CC -1- SUBUNIT: BINDS CYCLIN T1 (BY SIMILARITY).
CC -1- SUBCELLULAR LOCATION: NUCLEAR; NUCLEOLAR.
CC
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC
CC EMBL; K03454; AAA44322.1; -
CC HIV; K03454; TATSELI.
CC InterPro: IPR001831; HIV_Tat.
CC Pfam; PF00539; Tat; 1.
CC PRINTS; PR00055; HIVTATDOMAIN.
CC Transcription regulation; Activator; RNA-binding; Nuclear protein;
CC AIDS.
CC
CC SEQUENCE 99 AA; 11081 MW; B0ECD5199874A762 CRC64;

Query Match 69.6%; Score 384; DB 1; Length 99;
Best Local Similarity 72.1%; Pred. No. 1.8e-30;
Matches 62; Conservative 9; Mismatches 13; Indels 0; Gaps 0;

Qy 1 MEPVDPRLEPHKPGSQRTACTNCKCHCKCCYHCQVCFITKGLGISYGRKKRRRPPQ 60
Db 1 MDVDPNLEPHNPGSQRTPCNCKCHCKCCYHCQVCFITKGLGISYGRKKRRRPPQ 60
Qy 61 GSQTHQVLSKOPTSQSGKEPTGPK 86
Db 61 GQAHQVPIPEQPSQSRGDPGPK 86

RESULT 21
TAT_HV1U4
ID TAT_HV1U4 STANDARD; PRT; 101 AA.
AC P24738;
DT 01-MAR-1992 (Rel. 21, Created)
DT 01-MAR-1992 (Rel. 21, Last sequence update)
DT 16-OCT-2001 (Rel. 40, Last annotation update)
DE TAT protein (Transactivating regulatory protein).
GN TAT.
OS Human immunodeficiency virus type 1 (strain Ugandan / isolate U455)
OC Viruses; Retrovirdae; Lentivirus.
OX NCBI_TaxID=11703;
RN [1]
RP MEDLINE=91090981; PubMed=2265025;
RA Oram J.D., Downing R.G., Roff M., Clegg J.C.S., Serwadda D.,
RA Carswell J.W.;
RT "Nucleotide sequence of a Ugandan HIV-1 provirus reveals genetic
diversity from other HIV-1 isolates.";
RL AIDS Res. Hum. Retroviruses 6:1073-1078(1990).
CC -1- FUNCTION: TRANSCRIPTIONAL REGULATOR THAT ACTS BY BINDING TO THE
CC TRANS-ACTIVATING RESPONSE SEQUENCE (TAR) RNA ELEMENT AND
CC ACTIVATES TRANSCRIPTION INITIATION AND/OR ELONGATION FROM THE LTR
```

```
CC      PROMOTER.
CC      -!- SUBUNIT: BINDS CYCLIN T1 (BY SIMILARITY).
CC      -!- SUBCELLULAR LOCATION: NUCLEAR; NUCLEOLAR.
CC
CC      This SWISS-PROT entry is copyright. It is produced through a collaboration
CC      between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC      the European Bioinformatics Institute. There are no restrictions on its
CC      use by non-profit institutions as long as its content is in no way
CC      modified and this statement is not removed. Usage by and for commercial
CC      entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC      or send an email to license@isb-sib.ch).
CC
CC      EMBL; M62320; AAA75021.1; -.
CC      InterPro: IPR001831; HIV_Tat.
CC      Pfam; PF00539; Tat; 1.
CC      PRINTS; PR00055; HIVTATDOMAIN.
CC      Transcription regulation; Activator; RNA-binding; Nuclear protein;
CC      AIDS.
CC      SEQUENCE 101 AA; 11378 MW; 3C782A8055DAB249 CRC64;

Query Match      62.0%; Score 342; DB 1; Length 101;
Best Local Similarity 65.5%; Pred. No. 2e-26;
Matches 57; Conservative 11; Mismatches 19; Indels 0; Gaps 0;

QY 1 MEPVDPRLEPKHGPSQPKTACTNCYCKKCCFHCQVCFITAAALGISYGRKKRRRPPQ 60
DB 1 MEPVDNLEPKHGPSQPTTACNCYCKVCCWHCOLCFLKKGGLGISYGRKKRRRPPQ 60
QY 61 GSQTHQVSLSKQPTSQSGKEPTGPKET 87
DB 61 GSKDHQTLIPKQPLQSRVSAQGEES 87

RESULT 22
ID TAT_HV1B5 STANDARD; PRT; 58 AA.
AC P04612;
DT 13-AUG-1987 (Rel. 05, Created)
DT 13-AUG-1987 (Rel. 05, Last sequence update)
DT 16-OCT-2001 (Rel. 40, Last annotation update)
DE TAT protein (Transactivating regulatory protein) (Fragment).
GN TAT.
OS Human immunodeficiency virus type 1 (BH5 isolate) (HIV-1).
OC Viruses; Retroviral viruses; Retroviridae; Lentivirus.
OX NCBI_TaxID=11682;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=85111123; PubMed=2578615;
RA Ratner L., Haseltine W., Patarca R., Livak K.J., Starcich B.R.,
RA Josephs S.F., Doran E.R., Rafalski J.A., Whitehorn E.A.,
RA Baumeister K., Ivanoff L., Petteway S.R. Jr., Pearson M.L.,
RA Lautenberger J.A., Papas T.S., Graybe J., Chang N.T., Gallo R.C.,
RA Wong-Staal F.;
RT "Complete nucleotide sequence of the AIDS virus, HTLV-III.";
RL Nature 313:277-284(1985).
CC -!- FUNCTION: TRANSCRIPTIONAL REGULATOR THAT ACTS BY BINDING TO THE
CC TRANS-ACTIVATING RESPONSIVE SEQUENCE (TAR) RNA ELEMENT AND
CC ACTIVATES TRANSCRIPTION INITIATION AND/OR ELONGATION FROM THE LTR
CC PROMOTER.
CC -!- SUBUNIT: BINDS CYCLIN T1 (BY SIMILARITY).
CC -!- SUBCELLULAR LOCATION: NUCLEAR; NUCLEOLAR.
CC
CC      This SWISS-PROT entry is copyright. It is produced through a collaboration
CC      between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC      the European Bioinformatics Institute. There are no restrictions on its
CC      use by non-profit institutions as long as its content is in no way
CC      modified and this statement is not removed. Usage by and for commercial
CC      entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC      or send an email to license@isb-sib.ch).
CC
CC      EMBL; K02012; AAA44656.1; -.
CC      HIV; K02012; TAT5BH5.
```

```
DR      InterPro: IPR001831; HIV_Tat.
DR      Pfam; PF00539; Tat; 1.
DR      Transcription regulation; Activator; RNA-binding; Nuclear protein;
DR      AIDS.
DR      FT NON_TER 58
DR      SQ SEQUENCE 58 AA; 6800 MW; E36C21F8FFD813E3 CRC64;

Query Match      58.7%; Score 324; DB 1; Length 58;
Best Local Similarity 96.5%; Pred. No. 6.3e-25;
Matches 55; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 1 MEPVDPRLEPKHGPSQPKTACTNCYCKKCCFHCQVCFITAAALGISYGRKKRRR 57
DB 1 MEPVDPRLEPKHGPSQPKTACTTCYCKKCCFHCQVCFITAAALGISYGRKKRRR 57

RESULT 23
ID TAT_SIVCZ STANDARD; PRT; 100 AA.
AC P17285;
DT 01-AUG-1990 (Rel. 15, Created)
DT 01-AUG-1990 (Rel. 15, Last sequence update)
DT 16-OCT-2001 (Rel. 40, Last annotation update)
DE TAT protein (Transactivating regulatory protein).
GN TAT.
OS Chimpanzee immunodeficiency virus (SIV(cpz)) (CIV).
OC Viruses; Retroviral viruses; Retroviridae; Lentivirus.
OX NCBI_TaxID=11723;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=90259077; PubMed=2188136;
RA Huet T., Cheynier R., Meyerhans A., Roelants G., Wain-Hobson S.;
RT "Genetic organization of a chimpanzee lentivirus related to HIV-1.";
RL Nature 345:356-359(1990).
CC -!- FUNCTION: TRANSCRIPTIONAL REGULATOR THAT ACTS BY BINDING TO THE
CC TRANS-ACTIVATING RESPONSIVE SEQUENCE (TAR) RNA ELEMENT AND
CC ACTIVATES TRANSCRIPTION INITIATION AND/OR ELONGATION FROM THE LTR
CC PROMOTER.
CC -!- SUBUNIT: BINDS CYCLIN T1 (BY SIMILARITY).
CC -!- SUBCELLULAR LOCATION: NUCLEAR; NUCLEOLAR.
CC
CC      This SWISS-PROT entry is copyright. It is produced through a collaboration
CC      between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC      the European Bioinformatics Institute. There are no restrictions on its
CC      use by non-profit institutions as long as its content is in no way
CC      modified and this statement is not removed. Usage by and for commercial
CC      entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC      or send an email to license@isb-sib.ch).
CC
CC      EMBL; X52154; CAA36404.1; -.
CC      PIR; S09987; TNLJSI.
CC      HIV; X52154; TATSCP2.
CC      InterPro: IPR001831; HIV_Tat.
CC      Pfam; PF00539; Tat; 1.
CC      PRINTS; PR00055; HIVTATDOMAIN.
CC      Transcription regulation; Activator; RNA-binding; Nuclear protein;
CC      AIDS.
CC      SEQUENCE 100 AA; 11209 MW; 1B78830B90EED50E CRC64;

Query Match      57.2%; Score 315.5; DB 1; Length 100;
Best Local Similarity 60.7%; Pred. No. 6.7e-24;
Matches 54; Conservative 12; Mismatches 22; Indels 1; Gaps 1;

QY 1 MEPVDPRLEPKHGPSQPKTACTNCYCKKCCFHCQVCFITAAALGISYGRKKRRR-PP 59
DB 1 MDPIDPDLPEPKHGPSQPTVCNCCYKACCCYHCYCTKKGGLGISYGRKKRTTRRAP 60
QY 60 QGSOHQVSLSKQPTSQSGKEPTGPKETS 88
DB 61 AGSKNNQDSIPKQPLQSRGNKGESEKST 89
```

```

RESULT 24
TAT_HV2G1
ID TAT_HV2G1 STANDARD; PRT; 130 AA.
AC P18044;
DT 01-NOV-1990 (Rel. 16, Created)
DT 01-NOV-1990 (Rel. 16, Last sequence update)
DT 16-OCT-2001 (Rel. 40, Last annotation update)
DE TAT protein (Transactivating regulatory protein).
GN TAT.
OS Human immunodeficiency virus type 2 (isolate Ghana-1) (HIV-2).
OC Viruses; Retrovirdae; Lentivirus.
OX NCBI_TaxID=1171;
RN [1]
SEQUENCE FROM N.A.
MEDLINE=90122350; PubMed=2611042;
RA Hasegawa A., Tsujimoto H., Maki N., Ishikawa K.I., Miura T.,
RA Fukasawa M., Miki K., Hayami M.;
RT "Genomic divergence of HIV-2 from Ghana.";
RL AIDS Res. Hum. Retroviruses 5:593-604(1989).
CC -!- FUNCTION: TRANSCRIPTIONAL REGULATOR THAT ACTS BY BINDING TO THE
CC TRANS-ACTIVATING RESPONSIVE SEQUENCE (TAR) RNA ELEMENT AND
CC ACTIVATES TRANSCRIPTION INITIATION AND/OR ELONGATION FROM THE LTR
CC PROMOTER.
CC -!- SUBUNIT: BINDS CYCLIN T1 (BY SIMILARITY).
CC -!- SUBCELLULAR LOCATION: NUCLEAR; NUCLEOLAR.
CC
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC
CC EMBL; M30895; AAA43929.1; -
CC PIR; J50332; TNLJGG.
CC HIV; M30895; TAT$2GH1.
CC InterPro; IPR001831; HIV_Tat.
CC Pfam; PF00539; Tat; 1.
CC PRINTS; PR00055; HIVTATDOMAIN.
CC Transcription regulation; Activator; RNA-binding; Nuclear protein;
CC AIDS.
CC SEQUENCE 130 AA; 14580 MW; 95F308B537027D67 CRC64;
CC
CC Query Match 31.8%; Score 175.5; DB 1; Length 130;
CC Best Local Similarity 49.3%; Pred. No. 2.2e-10;
CC Matches 35; Conservative 9; Mismatches 24; Indels 3; Gaps 2;
CC
CC 17 OPKTACTN-CYCKKCCFHCQVCFFAALGISYGRKKRRRPPGQSTHOVSLSKQPTS 75
CC :| |||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
CC 45 RPLEACTNSCYCKQCSFHCQJCLFLKGLGIWYAKSR--RRTPRKTKTHSSASDKSIS 102
CC
CC 76 OSKGEPTGPK 86
CC :| |||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
CC 103 TRTGSQPTKE 113
CC
CC RESULT 25
TAT_HV2KR
ID TAT_HV2KR STANDARD; PRT; 130 AA.
AC Q74124;
DT 01-NOV-1997 (Rel. 35, Created)
DT 01-NOV-1997 (Rel. 35, Last sequence update)
DT 16-OCT-2001 (Rel. 40, Last annotation update)
DE TAT protein (Transactivating regulatory protein).
GN TAT.
OS Human immunodeficiency virus type 2 (isolate KF) (HIV-2).
OC Viruses; Retrovirdae; Lentivirus.
OX NCBI_TaxID=73484;
RN [1]
SEQUENCE FROM N.A.

```

```

RA Kraus G.K., Talbott R., Leavitt M., Luznick L., Schmidt A.,
RA Badel P., Bartz C., Morton W., Wong-Staal F., Looney D.J.;
RL Submitted (APR-1995) to the EMBL/GenBank/DBJ databases
CC -!- FUNCTION: TRANSCRIPTIONAL REGULATOR THAT ACTS BY BINDING TO THE
CC TRANS-ACTIVATING RESPONSIVE SEQUENCE (TAR) RNA ELEMENT AND
CC ACTIVATES TRANSCRIPTION INITIATION AND/OR ELONGATION FROM THE LTR
CC PROMOTER.
CC -!- SUBUNIT: BINDS CYCLIN T1 (BY SIMILARITY).
CC -!- SUBCELLULAR LOCATION: NUCLEAR; NUCLEOLAR.
CC
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC
CC EMBL; U22047; AAA64580.1; -
CC InterPro; IPR001831; HIV_Tat.
CC Pfam; PF00539; Tat; 1.
CC PRINTS; PR00055; HIVTATDOMAIN.
CC Transcription regulation; Activator; RNA-binding; Nuclear protein;
CC AIDS.
CC SEQUENCE 130 AA; 14440 MW; 2E328084003CC6E0 CRC64;
CC
CC Query Match 31.5%; Score 174; DB 1; Length 130;
CC Best Local Similarity 49.3%; Pred. No. 3e-10;
CC Matches 36; Conservative 8; Mismatches 25; Indels 4; Gaps 3;
CC
CC 17 OPKTACTN-CYCKKCCFHCQVCFFAALGISYGRKKRRRPPGQSTHOVSLSKQPTS 75
CC :| |||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
CC 45 QPLEACNPNCKKCCYHCQJCLFLKGLGICYERKGR--RRTPKKAHSSASDKSIS 102
CC
CC 76 OSKG-EPTGPKET 87
CC :| |||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
CC 103 TRTGSQTEKQKT 115
CC
CC RESULT 26
TAT_HV2ST
ID TAT_HV2ST STANDARD; PRT; 130 AA.
AC P20880;
DT 01-FEB-1991 (Rel. 17, Created)
DT 01-FEB-1991 (Rel. 17, Last sequence update)
DT 16-OCT-2001 (Rel. 40, Last annotation update)
DE TAT protein (Transactivating regulatory protein).
GN TAT.
OS Human immunodeficiency virus type 2 (isolate ST) (HIV-2).
OC Viruses; Retrovirdae; Lentivirus.
OX NCBI_TaxID=11721;
RN [1]
SEQUENCE FROM N.A.
MEDLINE=90112662; PubMed=2296086;
RA Kumar P., Hui H., Kappes J.C., Haggarty B.S., Hoxie J.A., Arya S.K.,
RA Shaw G.M., Hahn B.H.;
RT "Molecular characterization of an attenuated human immunodeficiency
RT virus type 2 isolate.";
RL J. Virol. 64:890-901(1990).
CC -!- FUNCTION: TRANSCRIPTIONAL REGULATOR THAT ACTS BY BINDING TO THE
CC TRANS-ACTIVATING RESPONSIVE SEQUENCE (TAR) RNA ELEMENT AND
CC ACTIVATES TRANSCRIPTION INITIATION AND/OR ELONGATION FROM THE LTR
CC PROMOTER.
CC -!- SUBUNIT: BINDS CYCLIN T1 (BY SIMILARITY).
CC -!- SUBCELLULAR LOCATION: NUCLEAR; NUCLEOLAR.
CC
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).

```



```

QY 71 -----KOPTOSK 78
Db 106 TRTRANSQAKKQK 118

RESULT 29
TAT_HV2D1
ID TAT_HV2D1 STANDARD; PRT; 130 AA.
AC P17759;
DT 01-AUG-1990 (Rel. 15, Created)
DT 01-AUG-1990 (Rel. 15, Last sequence update)
DT 01-NOV-1997 (Rel. 35, Last annotation update)
DE TAT protein (Transactivating regulatory protein).
GN TAT.
OS Human immunodeficiency virus type 2 (isolate D194) (HIV-2).
OC Viruses; Retrovirdae; Retroviridae; Lentivirus.
OX NCBI_TaxID=11713;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=89184631; PubMed=2467304;
RA Kuehnel H., von Briesen H., Dietrich U., Adamski M., Mix D.,
RA Blesert L., Kreutz R., Immanuel A., Henco K., Melchsnr C.,
RA Andresen R., Gelderblom H., Ruebsamen-Waigmann H.;
RT "Molecular cloning of two west African human immunodeficiency virus
RT type 2 isolates that replicate well in macrophages: a Gambian
RT isolate, from a patient with neurologic acquired immunodeficiency
RT syndrome, and a highly divergent Ghanaian isolate.";
RL Proc. Natl. Acad. Sci. U.S.A. 86:2383-2387(1989).
RN [2]
RP SEQUENCE FROM N.A.
RX MEDLINE=91045094; PubMed=2235509;
RA Kuehnel H., Kreutz R., Ruebsamen-Waigmann H.;
RT "Nucleotide sequence of HIV-2D194, an isolate from a Gambian case of
RT 'neuro-AIDS', which showed excellent growth in macrophages.";
RL Nucleic Acids Res. 18:6142-6142(1990).
CC -!- FUNCTION: TRANSCRIPTIONAL REGULATOR THAT ACTS BY BINDING TO THE
CC TRANS-ACTIVATING RESPONSIVE SEQUENCE (TAR) RNA ELEMENT AND
CC ACTIVATES TRANSCRIPTION INITIATION AND/OR ELONGATION FROM THE LTR
CC PROMOTER.
CC -!- SUBCELLULAR LOCATION: NUCLEAR; NUCLEOLAR.
CC -!- MISCELLANEOUS: THIS ISOLATE IS FROM A GAMBIAN CASE OF
CC 'NEURO-AIDS'.
CC
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC
CC EMBL; J04542; AAA76845.1; -
CC DR EMBL; X52223; CAA36469.1; -
CC DR PIR; S12157; S12157.
CC DR HIV; J04542; TAT52D194.
CC DR InterPro; IPR001831; HIV_Tat.
CC DR Pfam; PF00539; Tat; 1.
CC DR PRINTS; PR00055; HIVTATDOMAIN.
CC KW Transcription regulation; Activator; RNA-binding; Nuclear protein;
CC AIDS.
CC SEQUENCE 130 AA; 14597 MW; 1A659CB2F933C537 CRC64;

Query Match 29.2%; Score 161; DB 1; Length 130;
Best Local Similarity 33.9%; Pred. No. 5.3e-09;
Matches 39; Conservative 12; Mismatches 26; Indels 38; Gaps 5;

QY 2 EPVDRLEPWHKPGS-----QPKTACTN-CYCKKCCFHQV 36
Db 7 EP-ESLESYNEPSSCTSERDVTAGRAKQGBELLAQLRPLEACTNSCYCKQCSYHCQL 65
QY 37 CFITAAALGISYGRKKRRRRPPQSGSQTHVSL-----KQPTSQSKGEP 81

```

```

Db 66 CFLKGLGIWYARQGR--RRTPRKTKTHPPASPDKSISRTGDSQPTKKQKTP 118

RESULT 30
TAT_SIVM1
ID TAT_SIVM1 STANDARD; PRT; 130 AA.
AC P05911;
DT 01-NOV-1988 (Rel. 09, Created)
DT 01-NOV-1988 (Rel. 09, Last sequence update)
DT 16-OCT-2001 (Rel. 40, Last annotation update)
DE TAT protein (Transactivating regulatory protein).
GN TAT.
OS Simian immunodeficiency virus (Wm142-83 isolate) (SIV-MAC).
OC Viruses; Retrovirdae; Retroviridae; Lentivirus.
OX NCBI_TaxID=11733;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=87287230; PubMed=3649576;
RA Chakrabarti L., Guvader M., Alizon M., Daniel M.D., Destosiers R.C.,
RA Tiollais P., Sonigo P.;
RT "Sequence of simian immunodeficiency virus from macaque and its
RT relationship to other human and simian retroviruses.";
RL Nature 328:543-547(1987).
CC -!- FUNCTION: TRANSCRIPTIONAL REGULATOR THAT ACTS BY BINDING TO THE
CC TRANS-ACTIVATING RESPONSIVE SEQUENCE (TAR) RNA ELEMENT AND
CC ACTIVATES TRANSCRIPTION INITIATION AND/OR ELONGATION FROM THE LTR
CC PROMOTER.
CC -!- SUBUNIT: BINDS CYCLIN T1 (BY SIMILARITY).
CC -!- SUBCELLULAR LOCATION: NUCLEAR; NUCLEOLAR.
CC -!- MISCELLANEOUS: THIS IS A MACAQUE ISOLATE.
CC
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC
CC EMBL; Y00277; CAB46521.1; -
CC DR PIR; F28887; TNLJG3.
CC DR HIV; M16403; TAT5M142.
CC DR InterPro; IPR001831; HIV_Tat.
CC DR Pfam; PF00539; Tat; 1.
CC DR PRINTS; PR00055; HIVTATDOMAIN.
CC KW Transcription regulation; Activator; RNA-binding; Nuclear protein;
CC AIDS.
CC SEQUENCE 130 AA; 14545 MW; DA133BC4BAB7F521 CRC64;

Query Match 28.9%; Score 159.5; DB 1; Length 130;
Best Local Similarity 50.0%; Pred. No. 7.3e-09;
Matches 29; Conservative 10; Mismatches 16; Indels 3; Gaps 2;

QY 17 QPKTACTN-CYCKKCCFHQVCFITAAALGISYGRKKRRRRPPQSGSQTHVSLSKQP 73
Db 45 RPLEACYNTCYCKKCCYHCQFCFLKKGLGISY--EKSHRRRRTPKKAKANTSSASNEP 100

Search completed: August 26, 2002, 08:15:43
Job time: 347 sec

```

GenCore version 4.5  
Copyright (c) 1993 - 2000 Compugen Ltd.

OM protein - protein search, using sw model

Run on: August 26, 2002, 08:14:57 ; Search time 86.53 Seconds  
(without alignments)  
189.929 Million cell updates/sec

Title: US-09-509-239-23  
Perfect score: 552  
Sequence: 1 MEPVDRLEPWKHGSPQKT.....QSKGPTGPKETSGHHHHH 95

Scoring table: BLOSUM62  
Gapop 10.0 , Gapext 0.5

Searched: 562222 seqs, 172994929 residues

Total number of hits satisfying chosen parameters: 562222

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : SPTREMBL19.\*

- 1: sp\_archaea.\*
- 2: sp\_bacteria.\*
- 3: sp\_fungi.\*
- 4: sp\_human.\*
- 5: sp\_invertebrate.\*
- 6: sp\_mammal.\*
- 7: sp\_mmc.\*
- 8: sp\_organelle.\*
- 9: sp\_phage.\*
- 10: sp\_plant.\*
- 11: sp\_rodent.\*
- 12: sp\_virus.\*
- 13: sp\_vertebrate.\*
- 14: sp\_unclassified.\*
- 15: sp\_rvirus.\*
- 16: sp\_bacteriaph.\*
- 17: sp\_archaea.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

## SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	450	81.5	86	15	Q74087 human immun
2	448	81.2	86	15	Q900A7 human immun
3	446	80.8	86	15	Q99BV4 human immun
4	446	80.8	101	15	Q9DHA9 human immun
5	445	80.6	101	15	Q92880 human immun
6	443	80.3	86	15	Q99BV5 human immun
7	443	80.3	101	15	Q9DH20 human immun
8	442	80.1	86	15	O57293 simian-huma
9	441	79.9	101	15	Q99BV7 human immun
10	440	79.7	86	15	Q99BV6 human immun
11	440	79.7	86	15	Q99BV3 human immun
12	439	79.5	86	15	Q99BW1 human immun
13	439	79.5	86	15	Q99BW0 human immun
14	439	79.5	101	15	O92885 human immun
15	439	79.5	101	15	O92899 human immun
16	438	79.3	101	15	O56318 simian-huma

## ALIGNMENTS

## RESULT 1

Q74087 17 438 79.3 101 15 092893 human immun  
Q74087 18 438 79.3 101 15 099BV8 human immun  
AC Q74087 PRELIMINARY; PRT; 86 AA.  
DT 01-NOV-1996 (TREMBlrel. 01, Created)  
DT 01-NOV-1996 (TREMBlrel. 01, Last sequence update)  
DT 01-DEC-2001 (TREMBlrel. 19, Last annotation update)  
DE TAT PROTEIN (TRANSACTIVATING REGULATORY PROTEIN).  
GN TAT.  
OS Human immunodeficiency virus type 1.  
OC Viruses; Retrod viruses; Retroviridae; Lentivirus.  
OC NCBI\_TaxID=11676;  
RN [1]  
RP SEQUENCE FROM N.A.  
RC STRAIN=PM213;  
RX MEDLINE=90101366; PubMed=1688473;  
RA Cloyd M.W., Moore B.E.;  
RT "Spectrum of Biological Properties of Human Immunodeficiency Virus (HIV-1) Isolates."  
RL Virology 174:103-116(1990).  
RN [2]  
RP SEQUENCE FROM N.A.  
RC STRAIN=PM213;  
RA Iwatani Y.;  
RL Submitted (JUN-1996) to the EMBL/GenBank/DBJ databases.  
CC -1- FUNCTION: TRANSCRIPTIONAL REGULATOR THAT ACTS BY BINDING TO THE TRANS-ACTIVATING RESPONSIVE SEQUENCE (TAR) RNA ELEMENT AND ACTIVATES TRANSCRIPTION INITIATION AND/OR ELONGATION FROM THE LTR PROMOTER (BY SIMILARITY).  
CC EMBL: D86069; BAA13000.1; -.  
DR InterPro: IPR001831; HIV\_Tat.  
DR Pfam: PF00539; tat; 1.  
DR PRINTS: PR00055; HIVTATDOMAIN.  
DR KW Activator; Nuclear protein; RNA-binding; Transcription regulation.  
SQ SEQUENCE 86 AA; 9865 MW; 4DDC56CICE269115 CRC64;

Query Match 81.5%; Score 450; DB 15; Length 86;  
Best Local Similarity 91.9%; Pred. No. 2.2e-47;



```
Matches 79; Conservative 2; Mismatches 5; Indels 0; Gaps 0;

QY 1 MEPVDPRLEPKHGPSQPKTACTNCYCKCCFHCQVCFITAAALGISYGRKKRRRRPPQ 60
Db 1 MEPVDPRLEPKHGPSQPKTACTNCYCKCCFHCQVCFITAAALGISYGRKKRRRRRAHQ 60

QY 61 GSQTHOVSLSKOPTSQSKGPTGPKE 86
Db 61 NSQTHQVSLSKOPTSQSRGDPGTGPKE 86

RESULT 2
Q900A7 PRELIMINARY; PRT; 86 AA.
AC Q900A7;
DT 01-DEC-2001 (TReMBLrel. 19, Created)
DT 01-DEC-2001 (TReMBLrel. 19, Last sequence update)
DE TAT PROTEIN.
GN TAT.
OS Human immunodeficiency virus type 1.
OC Viruses; Retroid viruses; Retroviridae; Lentivirus.
OX NCBI_TaxID=11676;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=PL2.2;
RA Aldhous M.C.; Schwager M., Mok J.Y.Q., Yirell D.L., Leigh Brown A.J.,
RA Froebel K.S.;
RT "The natural history of pediatric HIV infection suggests an
RT association between the specificity of the cytotoxic T cells and
RT clinical outcome.";
RL Submitted (NOV-2000) to the EMBL/GenBank/DBJ databases.
CC -1- FUNCTION: TRANSCRIPTIONAL REGULATOR THAT ACTS BY BINDING TO THE
CC TRANS-ACTIVATING RESPONSE SEQUENCE (TAR) RNA ELEMENT AND
CC ACTIVATES TRANSCRIPTION INITIATION AND/OR ELONGATION FROM THE LTR
CC PROMOTER (BY SIMILARITY).
DR EMBL; AF324449; AAK14291.1; -
DR InterPro; IPR001831; HIV_Tat.
DR Pfam; PF00539; Tat; 1.
DR PRINTS; PRO0055; HIVTATDOMAIN.
DR Activator; Nuclear protein; RNA-binding; Transcription regulation.
SQ SEQUENCE 86 AA; 9842 MW; 60C7F0727C482F49 CRC64;

Query Match 80.8%; Score 446; DB 15; Length 86;
Best Local Similarity 89.5%; Pred. No. 6.8e-47;
Matches 77; Conservative 3; Mismatches 6; Indels 0; Gaps 0;

QY 1 MEPVDPRLEPKHGPSQPKTACTNCYCKCCFHCQVCFITAAALGISYGRKKRRRRPPQ 60
Db 1 MEPVDPRLEPKHGPSQPKTACTNCYCKCCFHCQVCFITAAALGISYGRKKRRRRSPQ 60

QY 61 GSQTHOVSLSKOPTSQSKGPTGPKE 86
Db 61 SSQTHQVSLSKOPTSQSRGDPGTGPKE 86

RESULT 4
Q9DHA9 PRELIMINARY; PRT; 101 AA.
AC Q9DHA9;
DT 01-MAR-2001 (TReMBLrel. 16, Created)
DT 01-MAR-2001 (TReMBLrel. 16, Last sequence update)
DT 01-DEC-2001 (TReMBLrel. 19, Last annotation update)
DE TAT PROTEIN (TRANSACTIVATING REGULATORY PROTEIN).
GN TAT.
OS Human immunodeficiency virus type 1.
OC Viruses; Retroid viruses; Retroviridae; Lentivirus.
OX NCBI_TaxID=11676;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=20457217; Pubmed=11000225;
RA Yuste E., Lopez-Galindez C., Domingo E.;
RT "Unusual Distribution of Mutations Associated with Serial Bottleneck
RT Passages of Human Immunodeficiency Virus Type 1.";
RL J. Virol. 74:9546-9552(2000)
CC -1- FUNCTION: TRANSCRIPTIONAL REGULATOR THAT ACTS BY BINDING TO THE
CC TRANS-ACTIVATING RESPONSE SEQUENCE (TAR) RNA ELEMENT AND
CC ACTIVATES TRANSCRIPTION INITIATION AND/OR ELONGATION FROM THE LTR
CC PROMOTER (BY SIMILARITY).
DR EMBL; AF256210; AAG16839.1; -
DR EMBL; AF256209; AAG16831.1; -
DR InterPro; IPR001831; HIV_Tat.
```

```
Matches 79; Conservative 2; Mismatches 5; Indels 0; Gaps 0;

QY 1 MEPVDPRLEPKHGPSQPKTACTNCYCKCCFHCQVCFITAAALGISYGRKKRRRRPPQ 60
Db 1 MEPVDPRLEPKHGPSQPKTACTNCYCKCCFHCQVCFITAAALGISYGRKKRRRRRAHQ 60

QY 61 GSQTHOVSLSKOPTSQSKGPTGPKE 86
Db 61 NSQTHQVSLSKOPTSQSRGDPGTGPKE 86

RESULT 2
Q900A7 PRELIMINARY; PRT; 86 AA.
AC Q900A7;
DT 01-DEC-2001 (TReMBLrel. 19, Created)
DT 01-DEC-2001 (TReMBLrel. 19, Last sequence update)
DE TAT PROTEIN.
GN TAT.
OS Human immunodeficiency virus type 1.
OC Viruses; Retroid viruses; Retroviridae; Lentivirus.
OX NCBI_TaxID=11676;
RN [1]
RP SEQUENCE FROM N.A.
RC MEDLINE=95074930; Pubmed=7983770;
RA Fang H., Pincus S.H.;
RA "Unique insertion sequence and pattern of CD4 expression in variants
RT selected with immunotoxins from human immunodeficiency virus type 1-
RT infected T cells.";
RL J. Virol. 69:75-81(1995).
RN [2]
RP SEQUENCE FROM N.A.
RA Fang H., Pincus S.H.;
RT "Spontaneous activation of human immunodeficiency virus type 1 in an
RT immunotoxin-resistant variant T cell line.";
RL Submitted (JUN-1998) to the EMBL/GenBank/DBJ databases.
RN [3]
RP SEQUENCE FROM N.A.
RC STRAIN=NL4-3;
RX MEDLINE=96036482; Pubmed=7483282;
RA Salminen M.O., Koch C., Sanders-Buell E., Ehrenberg P.K.,
RA Michael N.L., Carr J.K., Burke D.S., McCutchan F.E.;
RT "Recovery of virtually full-length HIV-1 provirus of diverse subtypes
RT from primary virus cultures using the polymerase chain reaction.";
RL Virology 213:80-86(1995)..
RN [4]
RP SEQUENCE FROM N.A.
RC STRAIN=NL4-3;
RX MEDLINE=86281827; Pubmed=3016298;
RA Adachi A., Gendelman H.E., Koenig S., Folks T., Willey R., Rabson A.,
RA Martin M.A.;
RT "Production of acquired immunodeficiency syndrome-associated
RT retrovirus in human and nonhuman cells transfected with an infectious
RT molecular clone.";
RL J. Virol. 59:284-291(1986).
DR EMBL; AF070521; AAC28449.1; -
DR EMBL; U26942; AAB60575.1; -
SQ SEQUENCE 86 AA; 9845 MW; 26B2281979769110 CRC64;

Query Match 81.2%; Score 448; DB 15; Length 86;
Best Local Similarity 90.7%; Pred. No. 3.9e-47;
Matches 78; Conservative 3; Mismatches 5; Indels 0; Gaps 0;

QY 1 MEPVDPRLEPKHGPSQPKTACTNCYCKCCFHCQVCFITAAALGISYGRKKRRRRPPQ 60
Db 1 MEPVDPRLEPKHGPSQPKTACTNCYCKCCFHCQVCFITAAALGISYGRKKRRRRRAHQ 60

QY 61 GSQTHOVSLSKOPTSQSKGPTGPKE 86
Db 61 NSQTHQVSLSKOPTSQSRGDPGTGPKE 86
```

```
DR Pfam: PF00539; Tat; 1.
DR PRINTS: PR00055: HIVTATDOMAIN.
KW Activator; Nuclear protein; RNA-binding; Transcription regulation.
SQ SEQUENCE 101 AA; 11425 MW; 1A1E5C93960630E CRC64;

Query Match      80.8%; Score 446; DB 15; Length 101;
Best Local Similarity 89.7%; Pred. No. 8e-47; 7; Indels 0; Gaps 0;
Matches 78; Conservative 2; Mismatches 0;

QY 1 MEPVDPRLPEPKHGPSQPKTACTNCYCKKCCFHCQVCFITAAALGISYGRKKRRRRPPQ 60
DB 1 MEPVDPRLPEPKHGPSQPKTACTNCYCKKCCFHCQVCFITAAALGISYGRKKRRRRPPQ 60
QY 61 GSQTHQVSLSKOPTSQSKGPTGPKE 87
DB 61 DSQTHQVSLSKOPTSQSKGPTGPKE 87

RESULT 5
Q92880 PRELIMINARY; PRT; 101 AA.
AC 092880;
DT 01-NOV-1998 (Tremblrel. 08, Created)
DT 01-NOV-1998 (Tremblrel. 08, Last sequence update)
DT 01-DEC-2001 (Tremblrel. 19, Last annotation update)
DE TAT PROTEIN (TRANSACTIVATING REGULATORY PROTEIN).
GN TAT.
OS Human immunodeficiency virus type 1.
OC Viruses; Retroid viruses; Retroviridae; Lentivirus.
OX NCBI_TaxID=11676;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=MBC200;
RX MEDLINE=20219736; PubMed=10754387;
RA Oelrichs R.B., Lawson V.A., Coates K.M., Chatfield C., Deacon N.J.,
RA McPhee D.A.;
RT "Rapid full-length genomic sequencing of two cytopathically
RT heterogeneous Australian primary HIV-1 isolates.";
RL J. Biomed. Sci. 7:128-135(2000).
RN [2]
RP SEQUENCE FROM N.A.
RC STRAIN=MBC200;
RA Oelrichs R.B., McPhee D.A., Deacon N.J.;
RT "Biological cloning and full-length sequence of two Australian HIV-1
RT primary clinical isolates with distinct cellular tropism and
RT cytopathicity.";
RL Submitted (JAN-1998) to the EMBL/GenBank/DBJ databases.
CC -!- FUNCTION: TRANSCRIPTIONAL REGULATOR THAT ACTS BY BINDING TO THE
CC TRANS-ACTIVATING RESPONSIVE SEQUENCE (TAR) RNA ELEMENT AND
CC ACTIVATES TRANSCRIPTION INITIATION AND/OR ELONGATION FROM THE LTR
CC PROMOTER (BY SIMILARITY).
DR EMBL: AF042100; AAD03194.1; -.
DR InterPro: IPR001831; HIV_Tat.
DR Pfam: PF00539; Tat; 1.
DR PRINTS: PR00055: HIVTATDOMAIN.
KW Activator; Nuclear protein; RNA-binding; Transcription regulation.
SQ SEQUENCE 101 AA; 11593 MW; A9F467BD68399235 CRC64;

Query Match      80.6%; Score 445; DB 15; Length 101;
Best Local Similarity 89.5%; Pred. No. 1.1e-46;
Matches 77; Conservative 3; Mismatches 6; Indels 0; Gaps 0;

QY 1 MEPVDPRLPEPKHGPSQPKTACTNCYCKKCCFHCQVCFITAAALGISYGRKKRRRRPPQ 60
DB 1 MEPVDPRLPEPKHGPSQPKTACTTCYCKKCCFHCQVCFITKGLGISYGRKKRRRRAPQ 60
QY 61 GSQTHQVSLSKOPTSQSKGPTGPKE 86
DB 61 DSETHQVSLSKOPTSQHRDPTGPKE 86
```

```
RESULT 6
Q99BV5 PRELIMINARY; PRT; 86 AA.
AC Q99BV5;
DT 01-JUN-2001 (Tremblrel. 17, Created)
DT 01-JUN-2001 (Tremblrel. 17, Last sequence update)
DT 01-DEC-2001 (Tremblrel. 19, Last annotation update)
DE TAT PROTEIN (TRANSACTIVATING REGULATORY PROTEIN).
GN TAT.
OS Human immunodeficiency virus type 1.
OC Viruses; Retroid viruses; Retroviridae; Lentivirus.
OX NCBI_TaxID=11676;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=PI2.1;
RA Aldhous M.C., Schwager M., Mok J.Y.Q., Virrell D.L., Leigh Brown A.J.,
RA Froebel K.S.;
RT "The natural history of pediatric HIV infection suggests an
RT association between the specificity of the cytotoxic T cells and
RT clinical outcome.";
RL Submitted (NOV-2000) to the EMBL/GenBank/DBJ databases.
CC -!- FUNCTION: TRANSCRIPTIONAL REGULATOR THAT ACTS BY BINDING TO THE
CC TRANS-ACTIVATING RESPONSIVE SEQUENCE (TAR) RNA ELEMENT AND
CC ACTIVATES TRANSCRIPTION INITIATION AND/OR ELONGATION FROM THE LTR
CC PROMOTER (BY SIMILARITY).
DR EMBL: AF324448; AAK14290.1; -.
DR InterPro: IPR001831; HIV_Tat.
DR Pfam: PF00539; Tat; 1.
DR PRINTS: PR00055: HIVTATDOMAIN.
KW Activator; Nuclear protein; RNA-binding; Transcription regulation.
SQ SEQUENCE 86 AA; 9828 MW; C1C6E1727D928438 CRC64;

Query Match      80.3%; Score 443; DB 15; Length 86;
Best Local Similarity 88.4%; Pred. No. 1.6e-46;
Matches 76; Conservative 4; Mismatches 6; Indels 0; Gaps 0;

QY 1 MEPVDPRLPEPKHGPSQPKTACTNCYCKKCCFHCQVCFITAAALGISYGRKKRRRRPPQ 60
DB 1 MDPVDPRLPEPKHGPSQPKTACTKCYCKKCCFHCQVCFITKGLGISYGRKKRRRRSPQ 60
QY 61 GSQTHQVSLSKOPTSQSKGPTGPKE 86
DB 61 SSQTHQVSLSKOPTSQPRGDPTGPKE 86

RESULT 7
Q9DH20 PRELIMINARY; PRT; 101 AA.
AC Q9DH20;
DT 01-MAR-2001 (Tremblrel. 16, Created)
DT 01-MAR-2001 (Tremblrel. 16, Last sequence update)
DT 01-DEC-2001 (Tremblrel. 19, Last annotation update)
DE TAT PROTEIN (TRANSACTIVATING REGULATORY PROTEIN).
GN TAT.
OS Human immunodeficiency virus type 1.
OC Viruses; Retroid viruses; Retroviridae; Lentivirus.
OX NCBI_TaxID=11676;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=20457217; PubMed=11000225;
RA Yuste E., Lopez-Galindez C., Domingo E.;
RT "Unusual Distribution of Mutations Associated with Serial Bottleneck
RT Passages of Human Immunodeficiency Virus Type 1.";
RL J. Virol. 74:9546-9552(2000).
RN [2]
RP SEQUENCE FROM N.A.
RA Yuste E., Lopez-Galindez C., Domingo E.;
RL Submitted (APR-2000) to the EMBL/GenBank/DBJ databases.
CC -!- FUNCTION: TRANSCRIPTIONAL REGULATOR THAT ACTS BY BINDING TO THE
CC TRANS-ACTIVATING RESPONSIVE SEQUENCE (TAR) RNA ELEMENT AND
CC ACTIVATES TRANSCRIPTION INITIATION AND/OR ELONGATION FROM THE LTR
CC PROMOTER (BY SIMILARITY).
```

```

DR EMBL: AF256211; AAG16848.1; -
DR EMBL: AF256204; AAG16788.1; -
DR EMBL: AF256205; AAG16797.1; -
DR EMBL: AF256206; AAG16805.1; -
DR EMBL: AF256207; AAG16814.1; -
DR EMBL: AF256208; AAG16822.1; -
DR InterPro: IPR001831; HIV_Tat.
DR Pfam: PF00539; Tat. 1.
DR PRINTS; PR00055; HIVTATDOMAIN.
KW Activator; Nuclear protein; RNA-binding; Transcription regulation.
SQ SEQUENCE 101 AA; 11453 MW; 831E576C7E82C649 CRC64;

Query Match 80.3%; Score 443; DB 15; Length 101;
Best Local Similarity 88.5%; Pred. No. 1.9e-46;
Matches 77; Conservative 3; Mismatches 7; Indels 0; Gaps 0;

QY 1 MEPVDRLPEPKHGPSQPKTACTNCYCKKCCFHCQVCFITAAAGISYGRKKRRRRPPQ 60
DB 1 MEPVDRLPEPKHGPSQPKTACTNCYCKKCCFHCQVCFITKGLGISYGRKKRRRRAPQ 60
QY 61 GSQTHQVSLSKOPTSQSKGPTGPKET 87
DB 61 DSQTHQVSLSKOPASQPGDPTGPKES 87

RESULT 8
O57293 PRELIMINARY; PRT; 86 AA.
ID O57293
AC O57293; 1998 (Tremblrel. 06, Created)
DT 01-JUN-1998 (Tremblrel. 06, Last sequence update)
DT 01-DEC-2001 (Tremblrel. 19, Last annotation update)
DE TAT PROTEIN (TRANSACTIVATING REGULATORY PROTEIN).
GN TAT.
OS Simian-Human immunodeficiency virus.
OC Viruses; Retroid viruses; Retroviridae; Lentivirus.
OX NCBI_TaxID=57667;
RN [1]
RP SEQUENCE FROM N.A.
RA Cayabyab M., Karlsson G.B., Etemad-Moghadam B., Hofmann W.,
RA Halloran M., Axthelms M.W., Letvin N.L., Sodroski J.G.;
RA Submitted (FEB-1998) to the EMBL/GenBank/DBJ databases.
RN [2]
RP SEQUENCE FROM N.A.
RC STRAIN-SHIV-4, AND HBXC2;
RX MEDLINE=92309177; PubMed=1613662;
RA Li J., Lord C.I., Haseltine W., Letvin N.L., Sodroski J.;
RT "Infection of cynomolgus monkeys with a chimeric HIV-1/SIVmac virus
RT that expresses the HIV-1 envelope glycoproteins."
RL J. Acquir. Immune Defic. Syndr. 5:639-646(1992).
RN [3]
RP SEQUENCE FROM N.A.
RC STRAIN-SHIV-4, AND HBXC2;
RA Reimann K.A., Li J.T., Voss G., Lekutis C., Tenner-Racz K., Racz P.,
RA Lin W., Montefiori D.C., Lee-Parritz D.E., Lu Y., Collman R.G.,
RA Sodroski J., Letvin N.L.;
RA Submitted (DEC-1997) to the EMBL/GenBank/DBJ databases.
CC -!- FUNCTION: TRANSCRIPTIONAL REGULATOR THAT ACTS BY BINDING TO THE
CC TRANS-ACTIVATING RESPONSIVE SEQUENCE (TAR) RNA ELEMENT AND
CC ACTIVATES TRANSCRIPTION INITIATION AND/OR ELONGATION FROM THE LTR
CC PROMOTER (BY SIMILARITY).
DR EMBL: AF041850; AAD21219.1; -
DR EMBL: AF038399; AAB99973.1; -
DR InterPro: IPR001831; HIV_Tat.
DR Pfam: PF00539; Tat. 1.
DR PRINTS; PR00055; HIVTATDOMAIN.
KW Activator; Nuclear protein; RNA-binding; Transcription regulation.
SQ SEQUENCE 86 AA; 9837 MW; 5132A4D9796364F5 CRC64;

Query Match 80.1%; Score 442; DB 15; Length 86;
Best Local Similarity 89.5%; Pred. No. 2.1e-46;

```

```

Matches 77; Conservative 3; Mismatches 6; Indels 0; Gaps 0;

QY 1 MEPVDRLPEPKHGPSQPKTACTNCYCKKCCFHCQVCFITAAAGISYGRKKRRRRPPQ 60
DB 1 MEPVDRLPEPKHGPSQPKTACTNCYCKKCCFHCQVCFITKGLGISYGRKKRRRRAPQ 60
QY 61 GSQTHQVSLSKOPTSQSKGPTGPKET 86
DB 61 NSQTHQASLSKOPTSQPRGDPGPKET 86

RESULT 9
Q99BV7 PRELIMINARY; PRT; 101 AA.
ID Q99BV7
AC Q99BV7; 2001 (Tremblrel. 17, Created)
DT 01-JUN-2001 (Tremblrel. 17, Last sequence update)
DT 01-DEC-2001 (Tremblrel. 19, Last annotation update)
DE TAT PROTEIN (TRANSACTIVATING REGULATORY PROTEIN).
GN TAT.
OS Human immunodeficiency virus type 1.
OC Viruses; Retroid viruses; Retroviridae; Lentivirus.
OX NCBI_TaxID=11676;
RN [1]
RP SEQUENCE FROM N.A.
RA Aldhous M.C., Schwager M., Mok J.Y.Q., Yirrell D.L., Leigh Brown A.J.;
RA "The natural history of pediatric HIV infection suggests an
RT association between the specificity of the cytotoxic T cells and
RT clinical outcome."
RL Submitted (NOV-2000) to the EMBL/GenBank/DBJ databases.
CC -!- FUNCTION: TRANSCRIPTIONAL REGULATOR THAT ACTS BY BINDING TO THE
CC TRANS-ACTIVATING RESPONSIVE SEQUENCE (TAR) RNA ELEMENT AND
CC ACTIVATES TRANSCRIPTION INITIATION AND/OR ELONGATION FROM THE LTR
CC PROMOTER (BY SIMILARITY).
DR EMBL: AF324446; AAK14288.1; -
DR InterPro: IPR001831; HIV_Tat.
DR Pfam: PF00539; Tat. 1.
DR PRINTS; PR00055; HIVTATDOMAIN.
KW Activator; Nuclear protein; RNA-binding; Transcription regulation.
SQ SEQUENCE 101 AA; 11517 MW; 03AB48CC28C9DE2 CRC64;

Query Match 79.9%; Score 441; DB 15; Length 101;
Best Local Similarity 88.5%; Pred. No. 3.3e-46;
Matches 77; Conservative 3; Mismatches 7; Indels 0; Gaps 0;

QY 1 MEPVDRLPEPKHGPSQPKTACTNCYCKKCCFHCQVCFITAAAGISYGRKKRRRRPPQ 60
DB 1 MEPVDRLPEPKHGPSQPKTACTPCYCKKCCFHCQVCFITKGLGISYGRKKRRRRAPQ 60
QY 61 GSQTHQVSLSKOPTSQSKGPTGPKET 87
DB 61 DSQTHQVSLSKOPTSQPRGDPGPKES 87

RESULT 10
Q99BV6 PRELIMINARY; PRT; 86 AA.
ID Q99BV6
AC Q99BV6; 2001 (Tremblrel. 17, Created)
DT 01-JUN-2001 (Tremblrel. 17, Last sequence update)
DT 01-DEC-2001 (Tremblrel. 19, Last annotation update)
DE TAT PROTEIN (TRANSACTIVATING REGULATORY PROTEIN).
GN TAT.
OS Human immunodeficiency virus type 1.
OC Viruses; Retroid viruses; Retroviridae; Lentivirus.
OX NCBI_TaxID=11676;
RN [1]
RP SEQUENCE FROM N.A.
RA Aldhous M.C., Schwager M., Mok J.Y.Q., Yirrell D.L., Leigh Brown A.J.;

```





RA Reimann K.A., Li J.T., Voss G., Lekutis C., Tenner-Racz K., Racz P.,  
 RA Lin W., Montefiori D.C., Lee-Parritz D.E., Lu Y., Collman R.G.,  
 RA Sodroski J., Letvin N.L.;  
 RL Submitted (DEC-1997) to the EMBL/GenBank/DBJ databases.  
 CC -!- FUNCTION: TRANSCRIPTIONAL REGULATOR THAT ACTS BY BINDING TO THE  
 CC TRANS-ACTIVATING RESPONSIVE SEQUENCE (TAR) RNA ELEMENT AND  
 CC ACTIVATES TRANSCRIPTION INITIATION AND/OR ELONGATION FROM THE LTR  
 CC PROMOTER (BY SIMILARITY).  
 DR EMBL: AF038398; AAB99963.1; -.  
 DR InterPro: IPR001831; HIV\_Tat.  
 DR Pfam: PF00539; Tat; 1.  
 DR PRINTS: PR00055; HIVTATDOMAIN.  
 KW Activator; Nuclear protein; RNA-binding; Transcription regulation.  
 SQ SEQUENCE 101 AA; 11629 MW; ACC85C209970DDE9 CRC64;

Query Match 79.3%; Score 438; DB 15; Length 101;  
 Best Local Similarity 88.4%; Pred. No. 7.6e-46;  
 Matches 76; Conservative 4; Mismatches 6; Indels 0; Gaps 0;

Qy 1 MEPVDPRLEPKHKGSPQKTCNCKKCFHCQVCFITAAALGISYGRKKRRRRPPQ 60  
 |||||  
 Db 1 MEPVDPRLEPKHKGSPQKTCNCKKCFHCQVCFITAAALGISYGRKKRRRRPPQ 60  
 |||||  
 Qy 61 GSQTHQVSLSKQPTSSQSGKEPTGPK 86  
 |||||  
 Db 61 NSQTHQASLSKQSPSGRGTGPK 86

RESULT 17  
 O92893 PRELIMINARY; PRT; 101 AA.  
 ID O92893  
 AC O92893;  
 DT 01-NOV-1998 (TrEMBLrel. 08, Created)  
 DT 01-NOV-1998 (TrEMBLrel. 08, Last sequence update)  
 DT 01-DEC-2001 (TrEMBLrel. 19, Last annotation update)  
 DE TAT PROTEIN (TRANSCRIPTING REGULATORY PROTEIN).  
 GN TAT.  
 OS Human immunodeficiency virus type 1.  
 OC Viruses; Retroid viruses; Retroviridae; Lentivirus.  
 OX NCBI\_TaxID=11676;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN=MBCC54;  
 RA Oelrichs R.B., McPhee D.A., Deacon N.J.;  
 RT "Genomic sequence of HIV-1 from four members of the Sydney Blood Bank  
 RT Cohort of long term non-progressors."  
 RL Submitted (JAN-1998) to the EMBL/GenBank/DBJ databases.  
 CC -!- FUNCTION: TRANSCRIPTIONAL REGULATOR THAT ACTS BY BINDING TO THE  
 CC TRANS-ACTIVATING RESPONSIVE SEQUENCE (TAR) RNA ELEMENT AND  
 CC ACTIVATES TRANSCRIPTION INITIATION AND/OR ELONGATION FROM THE LTR  
 CC PROMOTER (BY SIMILARITY).  
 DR EMBL: AF042103; AAD03220.1; -.  
 DR InterPro: IPR001831; HIV\_Tat.  
 DR Pfam: PF00539; Tat; 1.  
 DR PRINTS: PR00055; HIVTATDOMAIN.  
 KW Activator; Nuclear protein; RNA-binding; Transcription regulation.  
 SQ SEQUENCE 101 AA; 11509 MW; 974357617DBDEF30 CRC64;

Query Match 79.3%; Score 438; DB 15; Length 101;  
 Best Local Similarity 88.1%; Pred. No. 7.6e-46;  
 Matches 74; Conservative 6; Mismatches 7; Indels 0; Gaps 0;

Qy 1 MEPVDPRLEPKHKGSPQKTCNCKKCFHCQVCFITAAALGISYGRKKRRRRPPQ 60  
 |||||  
 Db 1 MEPVDPRLEPKHKGSPQKTCNCKKCFHCQVCFITAAALGISYGRKKRRRRPPQ 60  
 |||||  
 Qy 61 GSQTHQVSLSKQPTSSQSGKEPTGPK 87  
 |||||  
 Db 61 DSQTHQASLSKQPTSGRGTGPKES 87

RESULT 18  
 Q99BV8 PRELIMINARY; PRT; 101 AA.  
 ID Q99BV8  
 AC Q99BV8;  
 DT 01-JUN-2001 (TrEMBLrel. 17, Created)  
 DT 01-JUN-2001 (TrEMBLrel. 17, Last sequence update)  
 DT 01-DEC-2001 (TrEMBLrel. 19, Last annotation update)  
 DE TAT PROTEIN (TRANSCRIPTING REGULATORY PROTEIN).  
 GN TAT.  
 OS Human immunodeficiency virus type 1.  
 OC Viruses; Retroid viruses; Retroviridae; Lentivirus.  
 OX NCBI\_TaxID=11676;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN=P6.3;  
 RA Aldous M.C., Schwager M., Mok J.Y.Q., Yirell D.L., Leigh Brown A.J.,  
 RA Froebel K.S.;  
 RT "The natural history of pediatric HIV infection suggests an  
 RT association between the specificity of the cytotoxic T cells and  
 RT clinical outcome."  
 RL Submitted (NOV-2000) to the EMBL/GenBank/DBJ databases.  
 CC -!- FUNCTION: TRANSCRIPTIONAL REGULATOR THAT ACTS BY BINDING TO THE  
 CC TRANS-ACTIVATING RESPONSIVE SEQUENCE (TAR) RNA ELEMENT AND  
 CC ACTIVATES TRANSCRIPTION INITIATION AND/OR ELONGATION FROM THE LTR  
 CC PROMOTER (BY SIMILARITY).  
 DR EMBL: AF324444; AAK14287.1; -.  
 DR InterPro: IPR001831; HIV\_Tat.  
 DR Pfam: PF00539; Tat; 1.  
 DR PRINTS: PR00055; HIVTATDOMAIN.  
 KW Activator; Nuclear protein; RNA-binding; Transcription regulation.  
 SQ SEQUENCE 101 AA; 11615 MW; 5F80827F08100F03 CRC64;

Query Match 79.3%; Score 438; DB 15; Length 101;  
 Best Local Similarity 87.4%; Pred. No. 7.6e-46;  
 Matches 76; Conservative 4; Mismatches 7; Indels 0; Gaps 0;

Qy 1 MEPVDPRLEPKHKGSPQKTCNCKKCFHCQVCFITAAALGISYGRKKRRRRPPQ 60  
 |||||  
 Db 1 MEPVDPRLEPKHKGSPQKTCNCKKCFHCQVCFITAAALGISYGRKKRRRRPPQ 60  
 |||||  
 Qy 61 GSQTHQVSLSKQPTSSQSGKEPTGPK 87  
 |||||  
 Db 61 DSQTHQVSLSKQPTSGRGTGPKES 87

RESULT 19  
 Q9DQ29 PRELIMINARY; PRT; 101 AA.  
 ID Q9DQ29  
 AC Q9DQ29;  
 DT 01-MAR-2001 (TrEMBLrel. 16, Created)  
 DT 01-MAR-2001 (TrEMBLrel. 16, Last sequence update)  
 DT 01-DEC-2001 (TrEMBLrel. 19, Last annotation update)  
 DE TAT PROTEIN (TRANSCRIPTING REGULATORY PROTEIN).  
 GN TAT.  
 OS Human immunodeficiency virus type 1.  
 OC Viruses; Retroid viruses; Retroviridae; Lentivirus.  
 OX NCBI\_TaxID=11676;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN=WR27;  
 RA Salminen M.O., Ehrenberg P.K., Mascola J.R., Dayhoff D.E., Merling R.,  
 RA Blake B., Linder M., Hegerich S., Polonis V.R., Birx D.L., Robb M.L.,  
 RA McCutchan F.E., Michael N.L.;  
 RT "Construction and biological characterization of infectious molecular  
 RT clones of HIV-1 subtypes B and E (CRF01\_AE) generated by the  
 RT polymerase chain reaction."  
 RL Virology 278:103-110(2000).  
 RN [2]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN=WR27;  
 RA Salminen M.O., Ehrenberg P.K., Mascola J.R., Dayhoff D.E., Merling R.,

```
RA Blake B., Louder M., Hegerich S., Polonis V.R., Bix D.L., Robb M.L.,
RA McCutchan F.E., Michael N.L.;
RL Submitted (JUL-2000) to the EMBL/GenBank/DBJ databases.
CC -!- FUNCTION: TRANSCRIPTIONAL REGULATOR THAT ACTS BY BINDING TO THE
CC TRANS-ACTIVATING RESPONSIVE SEQUENCE (TAR) RNA ELEMENT AND
CC ACTIVATES TRANSCRIPTION INITIATION AND/OR ELONGATION FROM THE LTR
CC PROMOTER (BY SIMILARITY).
CC EMBL: AF286365; AAG30119.1; -.
DR InterPro: IPR001831; HIV_Tat.
DR Pfam: PF00539; Tat. 1.
DR PRINTS: PR00055; HIVTATDOMAIN.
KW Activator; Nuclear protein; RNA-binding; Transcription regulation.
SQ SEQUENCE 101 AA; 11572 MW; 41F97B5162E3C813 CRC64;

Query Match 79.2%; Score 437; DB 15; Length 101;
Best Local Similarity 86.2%; Pred. No. 1e-45;
Matches 75; Conservative 5; Mismatches 7; Indels 0; Gaps 0;

Qy 1 MEPVDPRLPEPKHPSQPKTACTNCYCKKCCFHCQVCFITAAAGISYGRKKRRRRPPQ 60
Db 1 MDPVDPRLPEPKHPSQPKTACTNCYCKKCCFHCQVCFITKGLGISYGRKKRRRRAPQ 60

Qy 61 GSQTHQVLSLKOPTSQSKGPTGPKET 87
Db 61 EGQTHQVLSLKOPTSQSKGPTGPKES 87

RESULT 20
ID Q902U5 PRELIMINARY; PRT; 101 AA.
AC Q902U5;
DT 01-DEC-2001 (TrEMBLrel. 19, Created)
DT 01-DEC-2001 (TrEMBLrel. 19, Last sequence update)
DT 01-DEC-2001 (TrEMBLrel. 19, Last annotation update)
DE TAT PROTEIN.
GN TAT.
OS Human immunodeficiency virus type 1.
OC Viruses; Retroid viruses; Retroviridae; Lentivirus.
OX NCBI_TaxID=11676;
RN [1]
RC STRAIN=ARCH014;
RA Carr J.K., Avila M., Gomez-Carrillo M., Salomon H., Hierholzer J.,
RA Watanaveeradej V., Pando M., Negrete M., Russell K.L., Sanchez J.,
RA Bix D.L., Andrade R., Vinholes J., McCutchan F.E.;
RT "Diverse BF Recombinants have spread widely since the introduction of
RT HIV-1 into South America.";
RL AIDS 0:0-0(2001).
DR EMBL: AY037266; AAL12635.1; -.
SQ SEQUENCE 101 AA; 11478 MW; 34432B8FD16F5DFC CRC64;

Query Match 79.2%; Score 437; DB 15; Length 101;
Best Local Similarity 87.4%; Pred. No. 1e-45;
Matches 76; Conservative 4; Mismatches 7; Indels 0; Gaps 0;

Qy 1 MEPVDPRLPEPKHPSQPKTACTNCYCKKCCFHCQVCFITAAAGISYGRKKRRRRPPQ 60
Db 1 MEPVDPRLPEPKHPSQPKTACTNCYCKKCCFHCQVCFITKGLGISYGRKKRRRRPPQ 60

Qy 61 GSQTHQVLSLKOPTSQSKGPTGPKET 87
Db 61 DSQTHQVLSLKOPTSQSKGPTGPKES 87

RESULT 21
ID O93022 PRELIMINARY; PRT; 101 AA.
AC O93022;
DT 01-NOV-1998 (TrEMBLrel. 08, Created)
DT 01-NOV-1998 (TrEMBLrel. 08, Last sequence update)
DT 01-DEC-2001 (TrEMBLrel. 19, Last annotation update)
```

```
DE TAT PROTEIN (TRANSACTIVATING REGULATORY PROTEIN).
GN TAT.
OS Human immunodeficiency virus type 1.
OC Viruses; Retroid viruses; Retroviridae; Lentivirus.
OX NCBI_TaxID=11676;
RN [1]
RC STRAIN=89ES061;
RA Lopez-galindez C.;
RL Submitted (MAY-1998) to the EMBL/GenBank/DBJ databases.
RN [2]
RC SEQUENCE FROM N.A.
RC STRAIN=89ES061;
RC MEDLINE=9905865; PubMed=9870318;
RA Olivares I., Casado C., Iglesias-Ussel M.D., Dietrich U.,
RA Lopez-galindez C.;
RT "Complete sequence of an infectious molecular clone derived from a
RT Spanish HIV type 1 isolate.";
RL AIDS Res. Hum. Retroviruses 14:1649-1651(1998).
CC -!- FUNCTION: TRANSCRIPTIONAL REGULATOR THAT ACTS BY BINDING TO THE
CC TRANS-ACTIVATING RESPONSIVE SEQUENCE (TAR) RNA ELEMENT AND
CC ACTIVATES TRANSCRIPTION INITIATION AND/OR ELONGATION FROM THE LTR
CC PROMOTER (BY SIMILARITY).
CC EMBL: AJ006287; CAA06950.1; -.
DR InterPro: IPR001831; HIV_Tat.
DR Pfam: PF00539; Tat. 1.
DR PRINTS: PR00055; HIVTATDOMAIN.
KW Activator; Nuclear protein; RNA-binding; Transcription regulation.
SQ SEQUENCE 101 AA; 11512 MW; 831E576C7CA2E449 CRC64;

Query Match 78.6%; Score 434; DB 15; Length 101;
Best Local Similarity 87.4%; Pred. No. 2.3e-45;
Matches 76; Conservative 3; Mismatches 8; Indels 0; Gaps 0;

Qy 1 MEPVDPRLPEPKHPSQPKTACTNCYCKKCCFHCQVCFITAAAGISYGRKKRRRRPPQ 60
Db 1 MEPVDPRLPEPKHPSQPKTACTNCYCKKCCFHCQVCFITKGLGISYGRKKRRRRAPQ 60

Qy 61 GSQTHQVLSLKOPTSQSKGPTGPKET 87
Db 61 DSQTHQVLSLKOPTSQSKGPTGPKES 87

RESULT 22
ID Q77688 PRELIMINARY; PRT; 101 AA.
AC Q77688;
DT 01-NOV-1996 (TrEMBLrel. 01, Created)
DT 01-NOV-1996 (TrEMBLrel. 01, Last sequence update)
DT 01-DEC-2001 (TrEMBLrel. 19, Last annotation update)
DE TAT PROTEIN (TRANSACTIVATING REGULATORY PROTEIN).
GN TAT.
OS Human immunodeficiency virus type 1.
OC Viruses; Retroid viruses; Retroviridae; Lentivirus.
OX NCBI_TaxID=11676;
RN [1]
RC STRAIN=CAMBRIDGE;
RA McIntosh A.A., Karpis A.;
RT "Nucleotide sequence of a Cambridge isolate of human immunodeficiency
RL virus type 1.";
RL Submitted (MAR-1992) to the EMBL/GenBank/DBJ databases.
CC -!- FUNCTION: TRANSCRIPTIONAL REGULATOR THAT ACTS BY BINDING TO THE
CC TRANS-ACTIVATING RESPONSIVE SEQUENCE (TAR) RNA ELEMENT AND
CC ACTIVATES TRANSCRIPTION INITIATION AND/OR ELONGATION FROM THE LTR
CC PROMOTER (BY SIMILARITY).
CC EMBL: D10112; BAA00996.1; -.
DR InterPro: IPR001831; HIV_Tat.
DR Pfam: PF00539; Tat. 1.
DR PRINTS: PR00055; HIVTATDOMAIN.
KW Activator; Nuclear protein; RNA-binding; Transcription regulation.
SQ SEQUENCE 101 AA; 11430 MW; B76CF7B13E60C3EF CRC64;
```







```

CC ACTIVATES TRANSCRIPTION INITIATION AND/OR ELONGATION FROM THE LTR
CC PROMOTER (BY SIMILARITY).
DR EMBL; M64491; AAA45079.1; -.
DR InterPro; IPR001831; HIV_Tat.
DR Pfam; PF00539; Tat; 1.
DR PRINTS; PR00055; HIVTATDOMAIN.
KW Activator; Nuclear protein; RNA-binding; Transcription regulation.
SQ SEQUENCE 101 AA; 11525 MW; 8B1A9AA50D4F658F CRC64;

Query Match 78.1%; Score 431; DB 15; Length 101;
Best Local Similarity 85.1%; Pred. No. 5.4e-45;
Matches 74; Conservative 4; Mismatches 9; Indels 0; Gaps 0;

QY 1 MEPVDPRLEPKPHGSGQPKTACTNCYCKKCCFHCQVCFTTAALGISYGRKKRRORRRPPQ 60
      |||||
DB 1 MEPVDPKLEPKPHGSGQPKTACNNCYCKKCCFHCQVCFTKKGLGISYGRKKRRORRRAPQ 60
      |||||

QY 61 GSQTHQVSLSKOPTSQSKGPTGPKEK 87
      |||||
DB 61 DSQTHQVSLSKQPASQPRGDTGPKEK 87
      |||||

```

GenCore version 4.5  
Copyright (c) 1993 - 2000 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: August 26, 2002, 08:08:31 ; Search time 99.16 Seconds  
(without alignments)  
338,284 Million cell updates/sec

Title: US-09-509-239-13  
Perfect score: 1688  
Sequence: 1 MGGKWSKSSVVGWPTVRM.....QSRGDTGPKETSGHHHHH 302

Scoring table: BLOSUM62  
Gapop 10.0 , Gapext 0.5

Searched: 747574 seqs, 11073796 residues

Total number of hits satisfying chosen parameters: 747574

Minimum DB seq length: 0  
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

Database : A\_Geneseq\_032802:.\*  
1: /SIDS1/gcgdata/hold-geneseq/geneseq-emb1/AA1980.DAT.\*  
2: /SIDS1/gcgdata/hold-geneseq/geneseq-emb1/AA1981.DAT.\*  
3: /SIDS1/gcgdata/hold-geneseq/geneseq-emb1/AA1982.DAT.\*  
4: /SIDS1/gcgdata/hold-geneseq/geneseq-emb1/AA1983.DAT.\*  
5: /SIDS1/gcgdata/hold-geneseq/geneseq-emb1/AA1984.DAT.\*  
6: /SIDS1/gcgdata/hold-geneseq/geneseq-emb1/AA1985.DAT.\*  
7: /SIDS1/gcgdata/hold-geneseq/geneseq-emb1/AA1986.DAT.\*  
8: /SIDS1/gcgdata/hold-geneseq/geneseq-emb1/AA1987.DAT.\*  
9: /SIDS1/gcgdata/hold-geneseq/geneseq-emb1/AA1988.DAT.\*  
10: /SIDS1/gcgdata/hold-geneseq/geneseq-emb1/AA1989.DAT.\*  
11: /SIDS1/gcgdata/hold-geneseq/geneseq-emb1/AA1990.DAT.\*  
12: /SIDS1/gcgdata/hold-geneseq/geneseq-emb1/AA1991.DAT.\*  
13: /SIDS1/gcgdata/hold-geneseq/geneseq-emb1/AA1992.DAT.\*  
14: /SIDS1/gcgdata/hold-geneseq/geneseq-emb1/AA1993.DAT.\*  
15: /SIDS1/gcgdata/hold-geneseq/geneseq-emb1/AA1994.DAT.\*  
16: /SIDS1/gcgdata/hold-geneseq/geneseq-emb1/AA1995.DAT.\*  
17: /SIDS1/gcgdata/hold-geneseq/geneseq-emb1/AA1996.DAT.\*  
18: /SIDS1/gcgdata/hold-geneseq/geneseq-emb1/AA1997.DAT.\*  
19: /SIDS1/gcgdata/hold-geneseq/geneseq-emb1/AA1998.DAT.\*  
20: /SIDS1/gcgdata/hold-geneseq/geneseq-emb1/AA1999.DAT.\*  
21: /SIDS1/gcgdata/hold-geneseq/geneseq-emb1/AA2000.DAT.\*  
22: /SIDS1/gcgdata/hold-geneseq/geneseq-emb1/AA2001.DAT.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	1688	100.0	302	20	AAY02351
2	1688	100.0	302	22	AAG63233
3	1688	100.0	411	22	AAG63235
4	1688	100.0	413	20	AAY02355
5	1688	100.0	413	22	AAG63237
6	1675	99.2	302	20	AAY02357
7	1675	99.2	302	22	AAG63239
8	1675	99.2	411	20	AAY02353
9	1139.5	67.5	215	20	AAY02349
10	1139.5	67.5	215	22	AAG63231
11	1139.5	67.5	324	20	AAY02352

12	1139.5	67.5	324	22	AAG63234
13	1139.5	67.5	326	20	AAY02354
14	1139.5	67.5	326	22	AAG63236
15	1120	66.4	206	21	AAY50795
16	1116	66.1	206	14	AAR38893
17	1115	66.1	206	21	AAB10054
18	1114	66.0	206	20	AAW89326
19	1111	65.8	206	7	AAP61515
20	1107	65.6	216	7	AAP60423
21	1092	64.7	206	20	AAW90179
22	1011	59.9	216	22	AAE04960
23	997	59.1	210	19	AAW53113
24	996	59.1	210	21	AAW77299
25	996	59.0	217	22	AAE04962
26	983	58.2	210	12	AAE12262
27	978	57.9	237	22	AAE04961
28	968	57.3	237	22	AAE04963
29	952	56.4	206	21	AAE04963
30	948.5	56.2	3025	22	AAE04963
31	934	55.3	206	21	AAE04963
32	927.5	54.9	206	9	AAE04963
33	910.5	53.9	207	21	AAE04963
34	903.5	53.5	207	11	AAE04963
35	901.5	53.4	207	21	AAE04963
36	895.5	53.1	219	21	AAE04963
37	895.5	52.9	217	21	AAE04963
38	893.5	52.9	217	21	AAE04963
39	885.5	52.5	207	21	AAE04963
40	878.5	52.0	207	21	AAE04963
41	867	51.4	206	21	AAE04963
42	860	50.9	206	21	AAE04963
43	834.5	49.4	209	19	AAW72998
44	831.5	49.3	209	9	AAE04963
45	796	47.2	212	19	AAW68481

ALIGNMENTS

RESULT 1  
AAY02351  
ID AAY02351 standard; Protein; 302 AA.  
XX AC AAY02351;  
XX DT 09-JUL-1999 (first entry)  
XX DE A representative HIV-1 Nef-Tat-His protein.  
XX KW HIV nef gene; fusion protein; HIV nef gene; Nef protein; Tat protein;  
XX KW vaccine; HIV infection; protein D.  
XX OS Synthetic.  
XX OS Human immunodeficiency virus type 1.  
XX PN WO9916884-A1.  
XX PD 08-APR-1999.  
XX PF 17-SEP-1998; 98WO-EP06040.  
XX PR 26-SEP-1997; 97GB-0020585.  
XX (SMIK ) SMITHKLINE BEECHAM BIOLOGICALS.  
XX Bruck C, Godart SAG, Marchand M;  
XX WPI; 1999-302282/25.  
XX N-PSDB; AAX35687.  
XX HIV Tat or Nef protein linked to a fusion partner  
XX Disclosure; Fig 2; 66pp; English.

Amino acid sequenc  
A representative L  
Amino acid sequenc  
Human Nef protein/  
Nef protein of HIV  
HIV-1 nef protein.  
HIV-1 nef protein  
Sequence of E' pro  
Sequence of LAV v1  
HIV-1 E' proteol  
HIV-1 jrl1 Nef pro  
Protein 6 containe  
HIV-1 (ATCC CRL 85  
HIV-1 jrl1 Nef (G2  
HIV-1 strain Oxi o  
Human tpa leader p  
Human tpa leader p  
HIV-1 non-subtype  
HIV-1 subtype C pr  
HIV-1 non-subtype  
Sequence encoded b  
HIV-1 non-subtype  
Sequence deduced f  
HIV-1 non-subtype  
HIV-1 non-subtype  
HIV-1 non-subtype  
HIV-1 non-subtype  
HIV-1 non-subtype  
HIV-1 non-subtype  
HIV-1 non-subtype  
HIV-1 non-subtype  
HIV isolate LAV,MA  
Sequence encoded b  
HIV-1 strain YBF30

XX The present sequence represents a representative HIV-1 Nef-Tat-His  
 CC protein. The protein is used in the creation of the fusion proteins of  
 CC the invention, in conjunction with a fusion partner (e.g. protein D).  
 CC The fusion protein can be used in a vaccine to prevent HIV infection.  
 XX  
 SQ Sequence 302 AA;

Query Match 100.0%; Score 1688; DB 20; Length 302;  
 Best Local Similarity 100.0%; Pred. No. 8.4e-156;  
 Matches 302; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MGKWSKSSVVGWPTVYRMRRAEPAADGVGAASRDLEKHGATTSNTAATNAACAWLEA 60  
 ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||  
 Db 1 mggkwsksavvgwptvremrrraepaadvgaasrdlekhgaitssntaatnaacawlea 60  
 ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||

QY 61 QEEEEVGFPTVQVPLRPMTYKAAVDLSHFLKEKGLGSLHSQRQDILDWIYHTQGY 120  
 ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||  
 Db 61 qeeeevgfptvpqplrpmtkaavdlshflkekggleghsqrqdlldwihtqgy 120  
 ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||

QY 121 FPDWQNTPGGVRYPLTFGWCKYKLVPEPKVVEEANKGENTSLHPVSLHGMDDPEREV 180  
 ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||  
 Db 121 fpdwqntpggvrypltfgwcyklvppepkveeankgentsllhpsvlhgmddperev 180  
 ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||

QY 181 LEWRFSRLAFHVARHRELHPEYFKNCTSEPVDPRLPEPKHPSQPKTACTNCYCKKCCFH 240  
 ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||  
 Db 181 lewrfsrlafhvarhrelhpeyfkntsepvdprrlepwhkpsqpkactactncycckccfh 240  
 ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||

QY 241 CQVCFITKALGISYGRKKRRRRPQSGQTHQVSLSKPTQSOSRQDTPGPKETSGHHHH 300  
 ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||  
 Db 241 cqvcfitkalgisygrkkrrrrppqsgqthqvsiskptsqsgdtpgpketsghhhh 300  
 ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||

QY 301 HH 302  
 ||  
 Db 301 hh 302

RESULT 2  
 AAG63233  
 ID AAG63233 standard; Protein; 302 AA.  
 AC AAG63233;  
 XX  
 XX 01-OCT-2001 (first entry)  
 XX  
 XX Amino acid sequence of a His-tagged Nef-Tat linked protein of HIV.  
 XX HIV; nef gene; tat gene; vaccine; gp120 gene; HIV viral load.  
 XX  
 XX Synthetic.  
 XX Human immunodeficiency virus.  
 XX  
 XX WO200154719-A2.  
 XX  
 XX 02-AUG-2001.  
 XX  
 XX 29-JAN-2001; 2001WO-EP00944.  
 XX  
 XX 31-JAN-2000; 2000GB-0002200.  
 XX 14-APR-2000; 2000GB-0009336.  
 XX 06-JUN-2000; 2000GB-0013806.  
 XX 28-JUN-2000; 2000WO-EP05998.  
 XX  
 XX (SMIK ) SMITHKLINE BEECHAM BIOLOGICALS.  
 XX  
 XX Voss G;  
 XX  
 XX WPI: 2001-476172/51.  
 XX N-PSDB: AAH42877.  
 XX  
 XX New use of a human immunodeficiency virus (HIV) Tat, or Nef, or Nef  
 PT linked to Tat (Nef-Tat) protein or polynucleotide and an HIV gp120

PT protein or polynucleotide for the manufacture of a vaccine -  
 XX Disclosure; Fig 1; 90pp; English.  
 XX  
 CC The present sequence represents a His-tagged Nef-Tat linked protein of  
 CC HIV. The protein is expressed in the yeast Pichia pastoris, and is used  
 CC to produce the vaccine of the invention. The specification describes  
 CC the use of HIV Tat, HIV Nef, or Nef-Tat; and HIV gp120 in the  
 CC manufacture of a vaccine. The vaccine is used for the prophylactic or  
 CC therapeutic immunization of humans against HIV. Tat, Nef or Nef-Tat act  
 CC in synergy with gp120 in the treatment and prevention of HIV. The  
 CC vaccine reduces the HIV viral load in HIV infected humans and results  
 CC in a maintenance of CD4+ levels over those levels found in the absence  
 CC of vaccination with HIV Tat, Nef or Nef-Tat and HIV gp120.  
 XX  
 SQ Sequence 302 AA;

Query Match 100.0%; Score 1688; DB 22; Length 302;  
 Best Local Similarity 100.0%; Pred. No. 8.4e-156;  
 Matches 302; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MGKWSKSSVVGWPTVYRMRRAEPAADGVGAASRDLEKHGATTSNTAATNAACAWLEA 60  
 ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||  
 Db 1 mggkwsksavvgwptvremrrraepaadvgaasrdlekhgaitssntaatnaacawlea 60  
 ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||

QY 61 QEEEEVGFPTVQVPLRPMTYKAAVDLSHFLKEKGLGSLHSQRQDILDWIYHTQGY 120  
 ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||  
 Db 61 qeeeevgfptvpqplrpmtkaavdlshflkekggleghsqrqdlldwihtqgy 120  
 ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||

QY 121 FPDWQNTPGGVRYPLTFGWCKYKLVPEPKVVEEANKGENTSLHPVSLHGMDDPEREV 180  
 ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||  
 Db 121 fpdwqntpggvrypltfgwcyklvppepkveeankgentsllhpsvlhgmddperev 180  
 ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||

QY 181 LEWRFSRLAFHVARHRELHPEYFKNCTSEPVDPRLPEPKHPSQPKTACTNCYCKKCCFH 240  
 ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||  
 Db 181 lewrfsrlafhvarhrelhpeyfkntsepvdprrlepwhkpsqpkactactncycckccfh 240  
 ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||

QY 241 CQVCFITKALGISYGRKKRRRRPQSGQTHQVSLSKPTQSOSRQDTPGPKETSGHHHH 300  
 ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||  
 Db 241 cqvcfitkalgisygrkkrrrrppqsgqthqvsiskptsqsgdtpgpketsghhhh 300  
 ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||

QY 301 HH 302  
 ||  
 Db 301 hh 302

RESULT 3  
 AAG63235  
 ID AAG63235 standard; Protein; 411 AA.  
 AC AAG63235;  
 XX  
 XX 01-OCT-2001 (first entry)  
 XX  
 XX Amino acid sequence of a His tagged LipD-Nef-Tat linked protein.  
 XX HIV; nef gene; tat gene; vaccine; gp120 gene; HIV viral load.  
 XX  
 XX Synthetic.  
 XX Human immunodeficiency virus.  
 XX  
 XX Key Location/Qualifiers  
 XX Peptide 1..109  
 XX /note= "ProtD fusion partner"  
 XX  
 XX WO200154719-A2.  
 XX  
 XX 02-AUG-2001.  
 XX  
 XX 29-JAN-2001; 2001WO-EP00944.  
 XX  
 XX 31-JAN-2000; 2000GB-0002200.  
 XX  
 XX WPI: 2001-476172/51.  
 XX N-PSDB: AAH42877.  
 XX  
 XX New use of a human immunodeficiency virus (HIV) Tat, or Nef, or Nef  
 PT linked to Tat (Nef-Tat) protein or polynucleotide and an HIV gp120

PR	14-APR-2000; 2000GB-0009336.
PR	06-JUN-2000; 2000GB-0013806.
PR	28-JUN-2000; 2000WO-EF05998.
XX	
XX	(SMIK ) SMITHKLINE BEECHAM BIOLOGICALS.
XX	
PI	Voss G;
XX	
DR	WPI; 2001-476172/51.
DR	N-PSDB; AAH42879.
XX	
PT	New use of a human immunodeficiency virus (HIV) Tat, or Nef, or Nef
PT	linked to Tat (Nef-Tat) protein or polynucleotide and an HIV gp120
PT	protein or polynucleotide for the manufacture of a vaccine -
XX	
XX	Disclosure; Fig 1; 90pp; English.
PS	
XX	
CC	The present sequence represents a His-tagged Nef-Tat linked protein of
CC	HIV, with a lipidation signal sequence (lipob) which is removed after
CC	processing and a ProT fusion partner. The protein is expressed in
CC	Escherichia coli, and is used to produce the vaccine of the invention.
CC	The specification describes the use of HIV Tat, HIV Nef, or Nef-Tat;
CC	and HIV gp120 in the manufacture of a vaccine. The vaccine is used for
CC	the prophylactic or therapeutic immunization of humans against HIV.
CC	Tat, Nef or Nef-Tat act in synergy with gp120 in the treatment and
CC	prevention of HIV. The vaccine reduces the HIV viral load in HIV
CC	infected humans and results in a maintenance of CD4+ levels over those
CC	levels found in the absence of vaccination with HIV Tat, Nef or Nef-Tat
CC	and HIV gp120.
XX	
XX	Sequence 411 AA;
SQ	

Query Match	100.0%;	Score 1688;	DB 22;	Length 411;
Best Local Similarity	100.0%;	Pred. No. 1.3e-155;		
Matches 302;	Conservative 0;	Mismatches 0;	Indels 0;	Gaps 0;
QY	1	MGGKWSXSSVGVPTVRMRRAEPAADGVGAASRDLEKHGCAITSSNTAATNAACAWLEA	60	
Db	110	mggkwsxssvgvptvrmmraepaadvgaasrdlekhgaitssntaataacawlea	169	
QY	61	QEEBEVGFVPTPOVPLRPMTYKAAVDLSHFLKEKGGLEGLIHSQRRORDILDWIYHTQGY	120	
Db	170	qeeevgfvtppvplrpmtykaavdlshflkekggleglihsqrqdlldwiyhtgy	229	
QY	121	FPDQWNTPGGVRYPITFGWCYKLVPEVPDKVBEANKGENTSLHPVSLHGMDDPEREV	180	
Db	230	fpdqwnypgvgvrypitfgwcyklvpvpdkveankgentsllhpvslhgmddperev	289	
QY	181	LEWRFDSRLAFHHVARELHPEYFNKCNSEPVDPLEPKWHPGSPKACTNCYCKKCCGFH	240	
Db	290	lewrfdsrlafhhvarelhpeyfnkcnsepvdplepkwkhpqspkactncyckkcgh	349	
QY	241	CQVCFIKALGISYGRKKRRORRPPGOSQTHOYVLSLKOPTSQSRGDDPTGPKETSGHHHH	300	
Db	350	cqvcfikaalgisvgrkrrqrrrrppqsgsqthqyslskqptsqsrqdpqpketsgghhh	409	
QY	301	HH 302		
Db	410	hh 411		

RESULT	4	
AAV02355		
ID	AAV02355 standard; Protein; 413 AA.	
XX		
AC	AAV02355;	
XX		
DT	09-JUL-1999 (first entry)	
XX		
DE	A representative LipD-Tat fusion protein.	
XX		
KW	HIV nef gene; fusion protein; HIV nef gene; Nef protein; Tat protein;	

KW	vaccine; HIV infection; protein D.
XX	
OS	Synthetic.
OS	Human immunodeficiency virus type 1.
XX	
XX	WO9916884-A1.
PN	
XX	
XX	08-APR-1999.
PD	
XX	
XX	17-SEP-1998; 98WO-EP06040.
PF	
XX	
XX	26-SEP-1997; 97GB-0020585.
PR	
XX	
XX	(SMIK ) SMITHKLINE BEECHAM BIOLOGICALS.
PA	
XX	
PI	Bruck C, Godart SAG, Marchand M;
XX	
DR	WPI; 1999-302282/25.
DR	N-PSDB; AAX35691.
XX	
XX	HIV Tat or Nef protein linked to a fusion partner
PT	
XX	
XX	Disclosure; Fig 2; 66pp; English.
PS	
XX	
CC	The present sequence represents a fusion protein comprising LipOD-HIV-1
CC	Tat. The protein is exemplifies the fusion proteins of
CC	the invention. The specification also describes fusion proteins
CC	comprising HIV-1 Nef protein. The fusion protein can be used in a
CC	vaccine to prevent HIV infection.
XX	
XX	Sequence 413 AA;
SO	

Query Match	100.0%;	Score 1688;	DB 20;	Length 413;
Best Local Similarity	100.0%;	Pred. NO. 1.3e-155;		
Matches 302;	Conservative 0;	Mismatches 0;	Indels 0;	Gaps 0;
Qy	1	MGGKWSKSSVVGVWTVTRMRRAEPAADGVCAASRDLEKHCAITSSNTAATNAACWLEA	60	
Db	112	mggkwsksvvgvwtvrmmraepaagdgvaasrdlekhgaitssntaataacwlea	171	
Qy	61	QEESEVGFPVTPQVPLRPMTYKAAVDLSHFLEKXGGLGLIHSQRRORDILDLYIHTQGY	120	
Db	172	qeeeevgfpvtqvpvlrptmykaavdlshflekkgleglihsqrrrdildlyhtqgy	231	
Qy	121	FPDQNYTPGPGVRYPLTFGCYKLVPEVPDKVEANKGENTSLHPVSLHGMDDPREV	180	
Db	232	fpdwqnytpgpgvrypltfgcykllvpepdkveeankgentsllhpvslhgmddperev	291	
Qy	181	LEWFDSERLAFHHVARELHDEYFNCTISEPVDPRLEPWKHPGSPKACTNICYCKKCCFH	240	
Db	292	lewifdsrlafhhvarelhpeyfnctisepvdprlepwhpgsqpktactnicyckkccfh	351	
Qy	241	COVCFITKALGISYGRKKRRRRPPGSGTOPTHYSLSKQPTSQSRGDPGPKETSGHHHH	300	
Db	352	cqvcfikalgisygrkkrrrrppgsgqchqvsllskpucsqsrgdptgpketsgghhhh	411	
Qy	301	HH 302		
Db	412	bb 413		

RESULT	5	
AAG63237		
ID	AAG63237 standard; Protein; 413 AA.	
XX		
XX	AAG63237;	
XX		
XX	01-OCT-2001 (first entry)	
DT		
XX		
DE	Amino acid sequence of a His tagged ProtD-Nef-Tat fusion protein.	
XX		
XX	HIV; nef gene; tat gene; vaccine; gp120 gene; HIV viral load.	
KW		

```

XX Synthetic.
OS Human immunodeficiency virus.
XX
XX FH Key Location/Qualifiers
XX Peptide 1..111
XX FT /note= "protD fusion partner"
XX
XX PN WO200154719-A2.
XX
XX PD 02-AUG-2001.
XX
XX PF 29-JAN-2001; 2001WO-EP00944.
XX
XX PR 31-JAN-2000; 2000GB-0002200.
XX PR 14-APR-2000; 2000GB-0009336.
XX PR 06-JUN-2000; 2000GB-0013806.
XX PR 28-JUN-2000; 2000WO-EP03998.
XX
XX PA (SMIK ) SMITHKLINE BEECHAM BIOLOGICALS.
XX
XX PI Voss G;
XX
XX DR WPI: 2001-476172/51.
XX DR N-PSDB; AAH42881.
XX
XX PT New use of a human immunodeficiency virus (HIV) Tat, or Nef, or Nef
XX linked to Tat (Nef-Tat) protein or polynucleotide and an HIV gp120
XX protein or polynucleotide for the manufacture of a vaccine -
XX
XX PS Disclosure; Fig 1; 90pp; English.
XX
XX CC The present sequence represents a His-tagged ProTD-Nef-Tat fusion
XX protein. The protein is expressed in Escherichia coli, and is used to
XX produce the vaccine of the invention. The specification describes
XX the use of HIV Tat, HIV Nef, or Nef-Tat; and HIV gp120 in the
XX manufacture of a vaccine. The vaccine is used for the prophylactic or
XX therapeutic immunization of humans against HIV. Tat, Nef or Nef-Tat act
XX in synergy with gp120 in the treatment and prevention of HIV. The
XX vaccine reduces the HIV viral load in HIV infected humans and results
XX in a maintenance of CD4+ levels over those levels found in the absence
XX of vaccination with HIV Tat, Nef or Nef-Tat and HIV gp120.
XX
XX SQ Sequence 413 AA;

Query Match 100.0%; Score 1688; DB 22; Length 413;
Best Local Similarity 100.0%; Pred. No. 1.3e-155;
Matches 302; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MGGKWSKSSVVGPTVRRMRRAEPADGVGAASRDLEKHGALTSSNTAATNAACAWLEA 60
Db 112 mggkwsksvvgptvrermraepaadvgaaardlekghaltsntaatnaacawlea 171
QY 61 QEEEEVGFPVTPQVPLRPMTYKAAVDLSHFLKEKGGLEGLHSORRODILDLYHTQGY 120
Db 172 qeeeevgfvtppqvlrptmcykaavdlshflkekggleghsqrddildlyhtcgy 231
QY 121 FPDQNYTPGPGVRYPLTFGCYKLVPEPDKVEEANKGENTSLHPVSLHGMDDPEREV 180
Db 232 fpdwnytpgpgvrypltfgcylkvpvdpdkveeankgentsllhpvslhgmddperev 291
QY 181 LEWRFDSRLAFHHVARELHPEYFNKCTSEPVDPRLPEWKPFGSQPKTACTNCYCKKCCFH 240
Db 292 lewrfdsrlafhhvarelhpeyfnkctsepvdprrlepwhkpgsqpktactncycckccfh 351
QY 241 CQVCFITKALGISYGRKKRRRPPGSGTHOVSLSKOPTSQSRGDPGTPKETSQHGH 300
Db 352 cqvcfitaalgisygrkkrrrrppgsgthqvslskoptsqsrqsgdpctgpketsghhh 411
QY 301 HH 302
Db 412 hh 413

```

```

RESULT 6
AAV02357
ID AAV02357 standard; Protein; 302 AA.
XX
AC AAV02357;
XX
DT 09-JUL-1999 (first entry)
XX
DE A representative HIV-1 mutant Tat-His protein.
XX
KW HIV nef gene; fusion protein; HIV nef gene; Nef protein; Tat protein;
KW vaccine; HIV infection; protein D.
XX
OS Synthetic.
OS Human immunodeficiency virus type 1.
XX
PN WO9916884-A1.
XX
PD 08-APR-1999.
XX
PF 17-SEP-1998; 98WO-EP06040.
XX
PR 26-SEP-1997; 97GB-0020585.
XX
PA (SMIK ) SMITHKLINE BEECHAM BIOLOGICALS.
XX
PI Bruck C, Godart SAG, Marchand M;
XX
DR WPI: 1999-302282/25.
DR N-PSDB; AAX35693.
XX
XX HIV Tat or Nef protein linked to a fusion partner
XX
XX Disclosure; Fig 2; 66pp; English.
XX
XX The present sequence represents a representative HIV-1 mutant Tat-His
XX protein. The protein is used in the creation of the fusion proteins of
XX the invention, in conjunction with a fusion partner (e.g. protein D).
XX The specification also describes fusion proteins, comprising HIV-1 Nef
XX protein. The fusion protein can be used in a vaccine to prevent HIV
XX infection.
XX
XX SQ Sequence 302 AA;

Query Match 99.2%; Score 1675; DB 20; Length 302;
Best Local Similarity 99.0%; Pred. No. 1.5e-154;
Matches 299; Conservative 2; Mismatches 1; Indels 0; Gaps 0;

QY 1 MGGKWSKSSVVGPTVRRMRRAEPADGVGAASRDLEKHGALTSSNTAATNAACAWLEA 60
Db 1 mggkwsksvvgptvrermraepaadvgaaardlekghaltsntaatnaacawlea 60
QY 61 QEEEEVGFPVTPQVPLRPMTYKAAVDLSHFLKEKGGLEGLHSORRODILDLYHTQGY 120
Db 61 qeeeevgfvtppqvlrptmcykaavdlshflkekggleghsqrddildlyhtcgy 120
QY 121 FPDQNYTPGPGVRYPLTFGCYKLVPEPDKVEEANKGENTSLHPVSLHGMDDPEREV 180
Db 121 fpdwnytpgpgvrypltfgcylkvpvdpdkveeankgentsllhpvslhgmddperev 180
QY 181 LEWRFDSRLAFHHVARELHPEYFNKCTSEPVDPRLPEWKPFGSQPKTACTNCYCKKCCFH 240
Db 181 lewrfdsrlafhhvarelhpeyfnkctsepvdprrlepwhkpgsqpktactncycckccfh 240
QY 241 CQVCFITKALGISYGRKKRRRPPGSGTHOVSLSKOPTSQSRGDPGTPKETSQHGH 300
Db 241 cqvcfitaalgisygrkkrrrrppgsgthqvslskoptsqsrqsgdpctgpketsghhh 300
QY 301 HH 302
Db 412 hh 413

```

Db 301 hh 302

RESULT 7

AAG63239

ID AAG63239 standard; Protein; 302 AA.

XX

AC AAG63239;

XX

DT 01-OCT-2001 (first entry)

XX

DE Amino acid sequence of a His-tagged mutant His protein of HIV.

XX

KW HIV; nef gene; tat gene; vaccine; gp120 gene; HIV viral load.

XX

OS Synthetic.

OS Human immunodeficiency virus.

XX

PN WO200154719-A2.

XX

PD 02-AUG-2001.

XX

PF 29-JAN-2001; 2001WO-EP00944.

XX

PR 31-JAN-2000; 2000GB-0002200.

PR 14-APR-2000; 2000GB-0009336.

PR 06-JUN-2000; 2000GB-0013806.

PR 28-JUN-2000; 2000WO-EP05998.

XX

PA (SMIK ) SMITHKLINE BEECHAM BIOLOGICALS.

XX

PI Voss G;

XX

DR WPI: 2001-476172/51.

DR N-PSDB; AAH42883.

XX

PT New use of a human immunodeficiency virus (HIV) Tat, or Nef, or Nef

PT linked to Tat (Nef-Tat) protein or polynucleotide and an HIV gp120

PT protein or polynucleotide for the manufacture of a vaccine .

XX

PS Disclosure; Fig 1; 90pp; English.

XX

CC The present sequence represents a His-tagged mutant His protein of HIV.

CC The protein is expressed in the yeast Pichia pastoris, and is used to

CC produce the vaccine of the invention. The specification describes

CC the use of HIV Tat, HIV Nef, or Nef-Tat; and HIV gp120 in the

CC manufacture of a vaccine. The vaccine is used for the prophylactic or

CC therapeutic immunization of humans against HIV. Tat, Nef or Nef-Tat act

CC in synergy with gp120 in the treatment and prevention of HIV. The

CC vaccine reduces the HIV viral load in HIV infected humans and results

CC in a maintenance of CD4+ levels over those levels found in the absence

CC of vaccination with HIV Tat, Nef or Nef-Tat and HIV gp120.

XX

SQ Sequence 302 AA;

Query Match 99.2%; Score 1675; DB 22; Length 302;

Best Local Similarity 99.0%; Pred. No. 1.5e-154;

Matches 299; Conservative 2; Mismatches 1; Indels 0; Gaps 0;

QY 1 MGSKWSSVVGWPTVRRMRRAEPAADGVGAASRDLEKHGALTSSNTAATNAACAWLEA 60

Db 1 mggkwsksvvgwptvrrmrtaepaadvgaasrdlekhgaltssntaataacawlea 60

QY 61 QEEEEVFPVTPQVPLRPMTYKAAVDLSHFLKEKGGLEGILHSQRQDILDLYHTQGY 120

Db 61 qeeeevfpvtpqvpplrmttykaavdlshflkekgglegilhsqrqgdildlyhtqgy 120

QY 121 FPDWQNTPGGVRYPLTGWCKYKLVVPDPKVEEANKGENTSLLHPVSLHGMDDPEREV 180

Db 121 fpdwqntpggvrypltgwcyklyvpdpdkveeankgentsllhpvslhgmddperev 180

QY 181 LEWRDLSRLAFHHVARELHPEYFKNCTSEFVDPRLPEFWKHPGQPKTACTNCTCYKCCCFH 240

Db 181 lewrdsrlafhhvarelhpeyfkntsefvpdrlefwkhpqsktactnctcykcccfh 240

QY 241 CQVCFITKALGISYGRKKRRRRPPQSGQTHQVLSUSKOPTSOSRGDPTGPKETSCHHHH 300

Db 241 cqvcfitaalgisygrkrrrrppqsgqthqvsiskgtskgeptgpketsghhh 300

QY 301 HH 302

Db 301 hh 302

RESULT 8

AAV02353

ID AAY02353 standard; Protein; 411 AA.

XX

AC AAY02353;

XX

DT 09-JUL-1999 (first entry)

XX

DE A representative Lipod-Tat-His fusion protein.

XX

KW HIV nef gene; fusion protein; HIV nef gene; Nef protein; Tat protein;

KW vaccine; HIV infection; protein D.

XX

OS Synthetic.

OS Human immunodeficiency virus type 1.

XX

PN WO9916884-A1.

XX

PD 08-APR-1999.

XX

PF 17-SEP-1998; 98WO-EP06040.

XX

PR 26-SEP-1997; 97GB-0020585.

XX

PA (SMIK ) SMITHKLINE BEECHAM BIOLOGICALS.

XX

PI Bruck C, Godart SAG, Marchand M;

XX

DR WPI: 1999-302282/25.

DR N-PSDB; AAX35689.

XX

PT HIV Tat or Nef protein linked to a fusion partner

XX

PS Disclosure; Fig 2; 66pp; English.

CC The present sequence represents a fusion protein comprising Lipod-HIV-1

CC Tat-His. The protein is exemplified the fusion proteins of

CC the invention. The specification also describes fusion proteins

CC comprising HIV-1 Nef protein. The fusion protein can be used in a

CC vaccine to prevent HIV infection.

XX

SQ Sequence 411 AA;

Query Match 99.2%; Score 1675; DB 20; Length 411;

Best Local Similarity 99.7%; Pred. No. 2.4e-154;

Matches 301; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 MGSKWSSVVGWPTVRRMRRAEPAADGVGAASRDLEKHGALTSSNTAATNAACAWLEA 60

Db 110 mggkwsksvvgwptvrrmrtaepaadvgaasrdlekhgaltssntaataacawlea 169

QY 61 QEEEEVFPVTPQVPLRPMTYKAAVDLSHFLKEKGGLEGILHSQRQDILDLYHTQGY 120

Db 170 qeeeevfpvtpqvpplrmttykaavdlshflkekgglegilhsqrqgdildlyhtqgy 229

QY 121 FPDWQNTPGGVRYPLTGWCKYKLVVPDPKVEEANKGENTSLLHPVSLHGMDDPEREV 180

Db 230 fpdwqntpggvrypltgwcyklyvpdpdkveeankgentsllhpvslhgmddperev 289

QY 181 LEWRDLSRLAFHHVARELHPEYFKNCTSEFVDPRLPEFWKHPGQPKTACTNCTCYKCCCFH 240

Db 290 lewrfdsrlafnhvarelhpeyfkntsepdprlepwkhpqsgpaktactnccyckccfh 349  
 |||||  
 Qy 241 CQVCFITKALGISYGRKKRRRRPPQGSOTHOVLSLKOPTSQSRGDPGPKETSGHHH 300  
 |||||  
 Db 350 CQVCFITKALGISYGRKKRRRRPPQGSOTHOVLSLKOPTSQSRGDPGPKETSGHHH 409  
 |||||  
 Qy 301 HH 302  
 ||  
 Db 410 hh 411

RESULT 9  
 AAY02349  
 ID AAY02349 standard; Protein; 215 AA.  
 XX  
 AC AAY02349;  
 XX  
 DT 09-JUL-1999 (first entry)  
 XX  
 DE A representative HIV-1 Nef-His protein.  
 XX  
 KW HIV nef gene; fusion protein; HIV nef gene; Nef protein; Tat protein;  
 KW vaccine; HIV infection; protein D.  
 XX  
 OS Synthetic.  
 OS Human immunodeficiency virus type 1.  
 XX  
 PN WO9916884-A1.  
 XX  
 PD 08-APR-1999.  
 XX  
 PF 17-SEP-1998; 98WO-EP06040.  
 XX  
 PR 26-SEP-1997; 97GB-0020585.  
 XX  
 PA (SMIK ) SMITHKLINE BEECHAM BIOLOGICALS.  
 XX  
 PI Bruck C, Godart SAG, Marchand M;  
 XX  
 DR WPI; 1999-302282/25.  
 DR N-PSDB; AAX35685.  
 XX  
 PT HIV Tat or Nef protein linked to a fusion partner  
 XX  
 PS Disclosure; Fig 2; 66pp; English.

CC The present sequence represents a representative HIV-1 Nef-His protein.  
 CC The protein is used in the creation of the fusion proteins of  
 CC the invention, in conjunction with a fusion partner (e.g. protein D).  
 CC The specification also describes fusion proteins comprising HIV-1 Tat  
 CC protein. The fusion protein can be used in a vaccine to prevent HIV  
 CC infection.  
 XX  
 SQ Sequence 215 AA;

Query Match 67.5%; Score 1139.5; DB 20; Length 215;  
 Best Local Similarity 71.2%; Pred. No. 1.2e-102;  
 Matches 215; Conservative 0; Mismatches 0; Indels 87; Gaps 1;

Qy 1 MGGKWSKSSVVGWPTVRMRRAEPAADGVGAASRDLEKHGATSSNTAATNAACAWLEA 60  
 |||||  
 Db 1 mggkwsksvvgwptvrermraepaadvgaasrdlekhgaltssntaataacawlea 60  
 |||||  
 Qy 61 QEEEEVGFPTPQVPLRPMYKAAVDLSHFLKEKGLHSGRRORODILDLWIYHTQGY 120  
 |||||  
 Db 61 qeeeevgfptpqvplrmpykaaavdlshflkekgleglhsgrrddldlwyhtqgy 120  
 |||||  
 Qy 121 FPDQNTPGGVRYPLTTFGWCYKLVPEPKVEANKGENTSLHHPVSLHGMDDPEREV 180  
 |||||  
 Db 121 fpdqnytpggvrypltfgwcyklvppepkveeankgentslhpsvshgmddperev 180  
 |||||

Qy 181 LEWRDSDRLAFNHVARELHPEYFKNTSEPDPRLEPWKHPQSQPKTACTNCCYCKCCFH 240  
 |||||  
 Db 181 lewrfdsrlafnhvarelhpeyfkntsepdprlepwkhpqsgpaktactnccyckccfh 206  
 |||||  
 Qy 241 CQVCFITKALGISYGRKKRRRRPPQGSOTHOVLSLKOPTSQSRGDPGPKETSGHHH 300  
 |||||  
 Db 207 -----tsghhhh 213  
 Qy 301 HH 302  
 ||  
 Db 214 hh 215

RESULT 10  
 AAG63231  
 ID AAG63231 standard; Protein; 215 AA.

XX  
 AC AAG63231;  
 XX  
 DT 01-OCT-2001 (first entry)  
 XX  
 DE Amino acid sequence of a His-tagged Nef protein of HIV.  
 XX  
 KW HIV; nef gene; tat gene; vaccine; gp120 gene; HIV viral load.  
 XX  
 OS Synthetic.  
 OS Human immunodeficiency virus.  
 XX  
 PN WO200154719-A2.  
 XX  
 PD 02-AUG-2001.

PF 29-JAN-2001; 2001WO-EP00944.  
 XX  
 PR 31-JAN-2000; 2000GB-0002200.  
 PR 14-APR-2000; 2000GB-0009336.  
 PR 06-JUN-2000; 2000GB-0013806.  
 PR 28-JUN-2000; 2000WO-EP03998.  
 XX  
 PA (SMIK ) SMITHKLINE BEECHAM BIOLOGICALS.

Voss G;

WPI; 2001-476172/51.  
 DR N-PSDB; AAH42875.

XX New use of a human immunodeficiency virus (HIV) Tat, or Nef, or Nef  
 XX linked to Tat (Nef-Tat) protein or polynucleotide and an HIV gp120  
 XX protein or polynucleotide for the manufacture of a vaccine -  
 PS Disclosure; Fig 1; 90pp; English.

CC The present sequence represents a His-tagged Nef protein of HIV. The  
 CC protein is expressed in the yeast Pichia pastoris, and is used to  
 CC produce the vaccine of the invention. The specification describes  
 CC the use of HIV Tat, HIV Nef, or Nef-Tat; and HIV gp120 in the  
 CC manufacture of a vaccine. The vaccine is used for the prophylactic or  
 CC therapeutic immunization of humans against HIV. Tat, Nef or Nef-Tat act  
 CC in synergy with gp120 in the treatment and prevention of HIV. The  
 CC vaccine reduces the HIV viral load in HIV infected humans and results  
 CC in a maintenance of CD4+ levels over those levels found in the absence  
 CC of vaccination with HIV Tat, Nef or Nef-Tat and HIV gp120.

SQ Sequence 215 AA;

Query Match 67.5%; Score 1139.5; DB 22; Length 215;  
 Best Local Similarity 71.2%; Pred. No. 1.2e-102;  
 Matches 215; Conservative 0; Mismatches 0; Indels 87; Gaps 1;

Qy 1 MGGKWSKSSVVGWPTVRMRRAEPAADGVGAASRDLEKHGATSSNTAATNAACAWLEA 60  
 |||||  
 Db 1 mggkwsksvvgwptvrermraepaadvgaasrdlekhgaltssntaataacawlea 60  
 |||||



```

QY 61 QEEVEGFPVTPQVPLRPMTYKAADVLSHFLKEKGLGSLHSQRQDILDLYHTQGY 120
Db 61 qeevevgfvtppqvprrpmtykaadvlsflkkggleglinsrrrqdildlyhtqgy 120
QY 121 FPDWQNYTPGPGVRYPLTFGWCKYKLVPEPDKVEEANKGENTSLHPVSLHGMDDPEREV 180
Db 121 fpdwqnytpgpgvrypltfgwckylvpepdkveeankgentsllhpvsilhgmddperev 180
QY 181 LEWRFSRLAFHHVARELHPEYFKNCTSEPDVRLPEPKHPSQPKTACTNVCYKCKCFH 240
Db 181 lewrfsrlafhhvarelhpeyfknc----- 206
QY 241 CQVCFTTKALGISYGRKKRRRPPQSGQTHQVSLSKQPTQSOSRGDPTGPKETSGHHH 300
Db 207 -----tsghhhh 213
QY 301 HH 302
Db 214 hh 215

RESULT 11
AAY02352
ID AAY02352 standard; Protein; 324 AA.
XX
AC AAY02352;
XX
DT 09-JUL-1999 (first entry)
XX
DE A representative LipD-Nef-His fusion protein.
XX
KW HIV nef gene; fusion protein; HIV nef gene; Nef protein; Tat protein;
KW vaccine; HIV infection; protein D.
XX
OS Synthetic.
OS Human Immunodeficiency virus type 1.
XX
PN WO9916884-A1.
XX
PD 08-APR-1999.
XX
PF 17-SEP-1998; 98WO-EP06040.
XX
PR 26-SEP-1997; 97GB-0020585.
XX
PA (SMIK ) SMITHKLINE BEECHAM BIOLOGICALS.
XX
PI Bruck C, Godart SAG, Marchand M;
XX
DR WPI; 1999-302282/25.
XX
N-PSDB; AAX35688.
XX
PT HIV Tat or Nef protein linked to a fusion partner
XX
PS Disclosure; Fig 2; 66pp; English.
XX
CC The present sequence represents a fusion protein comprising LipD-HIV-1
CC Nef-His. The protein is exemplified the fusion proteins of
CC the invention. The specification also describes fusion proteins
CC comprising HIV-1 Tat protein. The fusion protein can be used in a
CC vaccine to prevent HIV infection.
XX
SQ Sequence 324 AA;

Query Match 67.5%; Score 1139.5; DB 20; Length 324;
Best Local Similarity 71.2%; Pred. No. 2-2e-102;
Matches 215; Conservative 0; Mismatches 0; Indels 87; Gaps 1;

QY 1 MGKWSKSSVGVNPTVRRMRRAEPADGVGAASRDLEKHGATTSNTAATNAACAWLEA 60
Db 110 mggkwsksvvgvntvrrmrreapadvggaasrdlekbgaltssntaataacawlea 169

```

```

QY 61 QEEVEGFPVTPQVPLRPMTYKAADVLSHFLKEKGLGSLHSQRQDILDLYHTQGY 120
Db 170 qeevevgfvtppqvprrpmtykaadvlsflkkggleglinsrrrqdildlyhtqgy 229
QY 121 FPDWQNYTPGPGVRYPLTFGWCKYKLVPEPDKVEEANKGENTSLHPVSLHGMDDPEREV 180
Db 230 fpdwqnytpgpgvrypltfgwckylvpepdkveeankgentsllhpvsilhgmddperev 289
QY 181 LEWRFSRLAFHHVARELHPEYFKNCTSEPDVRLPEPKHPSQPKTACTNVCYKCKCFH 240
Db 290 lewrfsrlafhhvarelhpeyfknc----- 315
QY 241 CQVCFTTKALGISYGRKKRRRPPQSGQTHQVSLSKQPTQSOSRGDPTGPKETSGHHH 300
Db 316 -----tsghhhh 322
QY 301 HH 302
Db 323 hh 324

RESULT 12
AAG63234
ID AAG63234 standard; Protein; 324 AA.
XX
AC AAG63234;
XX
DT 01-OCT-2001 (first entry)
XX
DE Amino acid sequence of a His tagged LipD-Nef of HIV.
XX
KW HIV; nef gene; tat gene; vaccine; gp120 gene; HIV viral load.
XX
OS Synthetic.
OS Human Immunodeficiency virus.
XX
FH Key Location/Qualifiers
FT Peptide 1..109
FT /note= "ProtD fusion partner"
XX
PN WO200154719-A2.
XX
PD 02-AUG-2001.
XX
PF 29-JAN-2001; 2001WO-EP00944.
XX
PR 31-JAN-2000; 2000GB-0002200.
PR 14-APR-2000; 2000GB-0009336.
PR 06-JUN-2000; 2000GB-0013806.
PR 28-JUN-2000; 2000WO-EP05998.
XX
PA (SMIK ) SMITHKLINE BEECHAM BIOLOGICALS.
XX
PI Voss G;
XX
DR WPI; 2001-476172/51.
DR N-PSDB; AAH42878.
XX
PT New use of a human immunodeficiency virus (HIV) Tat, or Nef, or Nef
PT linked to Tat (Nef-Tat) protein or polynucleotide and an HIV gp120
PT protein or polynucleotide for the manufacture of a vaccine.
XX
PS Disclosure; Fig 1; 90pp; English.
XX
CC The present sequence represents a His-tagged Nef protein of HIV, with
CC a lipidation signal sequence (LipD) which is removed after processing.
CC The protein is expressed in Escherichia coli, and is used to
CC produce the vaccine of the invention. The specification describes
CC the use of HIV Tat, HIV Nef, or Nef-Tat; and HIV gp120 in the
CC manufacture of a vaccine. The vaccine is used for the prophylactic or
CC therapeutic immunization of humans against HIV. Tat, Nef or Nef-Tat act
CC in synergy with gp120 in the treatment and prevention of HIV. The

```

CC vaccine reduces the HIV viral load in HIV infected humans and results  
CC in a maintenance of CD4+ levels over those levels found in the absence  
CC of vaccination with HIV Tat, Nef or Nef-Tat and HIV gp120.

XX  
SQ Sequence 324 AA;

Query Match 67.5%; Score 1139.5; DB 22; Length 324;  
Best Local Similarity 71.2%; Pred. No. 2.2e-102;  
Matches 215; Conservative 0; Mismatches 0; Indels 87; Gaps 1;

QY 1 MGCKSKSSVVGPTVRRMRRAEPAADGVGAASRDLEKHGATTSNTAATNAACAWLEA 60  
DB 110 mgdkwkskssvvgptvrrmrreapaaadvgvaasrdlekhgaitssntaatnaacawlea 169  
QY 61 QEEEEVGFPVTPVPLRPMTYKAAVDLSHFLKEKGLGSLHSORRQDILDWYHTQGY 120  
DB 170 qeeeevgfvtppvplrpmtykaavdlshflkekgglegllhsqrrqdlldwlyhtqgy 229  
QY 121 FPDQMYTTPGQVRYPLTFGWCYKLVPEPDKVEEANKGENTSLLHPVSLHGMDPEREV 180  
DB 230 fpdqmytppgvyrypltfwgcyklypvepdkveeankgentsllhpvslhgmdperv 289  
QY 181 LEWRFSRLAFHHVARELHPEYFKNCTSEVPDRLEPWHKPGSQPKTACTNCYCKKCCFH 240  
DB 290 lewrfsrlafhhvarelhpeyfknc----- 315  
QY 241 CQVCFITKALGISYGRKKRRRRPQGSOTHOVSLSKPTQSQRGDPGTGPKETSGHHH 300  
DB 316 -----tsghhhh 322

QY 301 HH 302  
DB 323 hh 324

## RESULT 13

AA02334  
ID AAY02354 standard; Protein; 326 AA.

XX  
AC AAY02354;

DT 09-JUL-1999 (first entry)

XX A representative Lipod-Nef fusion protein.

DE HIV nef gene; fusion protein; HIV nef gene; Nef protein; Tat protein;  
KW vaccine; HIV infection; protein D.

XX  
OS Synthetic.

OS Human immunodeficiency virus type 1.

XX WO9916884-A1.

XX 08-APR-1999.

XX 17-SEP-1998; 98WO-EP06040.

XX 26-SEP-1997; 97GB-0020585.

XX (SMIK ) SMITHKLINE BEECHAM BIOLOGICALS.

XX Bruck C, Godart SAG, Marchand M;

XX WPI; 1999-302282/25.

XX N-PSDB; AAX35690.

XX HIV Tat or Nef protein linked to a fusion partner

XX Disclosure; Fig 2; 66pp; English.

XX The present sequence represents a fusion protein comprising Lipod-HIV-1  
CC Nef. The protein is exemplified by the fusion proteins of

CC the invention. The specification also describes fusion proteins  
CC comprising HIV-1 Tat protein. The fusion protein can be used in a  
XX vaccine to prevent HIV infection.

SQ Sequence 326 AA;

Query Match 67.5%; Score 1139.5; DB 20; Length 326;  
Best Local Similarity 71.2%; Pred. No. 2.2e-102;  
Matches 215; Conservative 0; Mismatches 0; Indels 87; Gaps 1;

QY 1 MGCKSKSSVVGPTVRRMRRAEPAADGVGAASRDLEKHGATTSNTAATNAACAWLEA 60  
DB 112 mgdkwkskssvvgptvrrmrreapaaadvgvaasrdlekhgaitssntaatnaacawlea 171  
QY 61 QEEEEVGFPVTPVPLRPMTYKAAVDLSHFLKEKGLGSLHSORRQDILDWYHTQGY 120  
DB 172 qeeeevgfvtppvplrpmtykaavdlshflkekgglegllhsqrrqdlldwlyhtqgy 231  
QY 121 FPDQMYTTPGQVRYPLTFGWCYKLVPEPDKVEEANKGENTSLLHPVSLHGMDPEREV 180  
DB 232 fpdqmytppgvyrypltfwgcyklypvepdkveeankgentsllhpvslhgmdperv 291  
QY 181 LEWRFSRLAFHHVARELHPEYFKNCTSEVPDRLEPWHKPGSQPKTACTNCYCKKCCFH 240  
DB 292 lewrfsrlafhhvarelhpeyfknc----- 317  
QY 241 CQVCFITKALGISYGRKKRRRRPQGSOTHOVSLSKPTQSQRGDPGTGPKETSGHHH 300  
DB 318 -----tsghhhh 324

QY 301 HH 302  
DB 325 hh 326

## RESULT 14

AA063236  
ID AAG63236 standard; Protein; 326 AA.

XX  
AC AAG63236;

DT 01-OCT-2001 (first entry)

DE Amino acid sequence of a His tagged ProtD-Nef fusion protein.

XX HIV; nef gene; tat gene; vaccine; gp120 gene; HIV viral load.

XX  
OS Synthetic.

OS Human immunodeficiency virus.

XX Key Location/Qualifiers

FT Peptide 1..111

FT /note- "ProtD fusion partner"

XX WO200154719-A2.

XX 02-AUG-2001.

XX 29-JAN-2001; 2001WO-EP00944.

XX 31-JAN-2000; 2000GB-0002200.

XX 14-APR-2000; 2000GB-0009336.

XX 06-JUN-2000; 2000GB-0013806.

XX 28-JUN-2000; 2000WO-EP05998.

XX (SMIK ) SMITHKLINE BEECHAM BIOLOGICALS.

XX Voss G;

XX WPI; 2001-476172/51.

XX N-PSDB; AAH42880.

PT New use of a human immunodeficiency virus (HIV) Tat, or Nef, or Nef  
PT linked to Tat (Nef-Tat) protein or polynucleotide and an HIV gp120  
PT protein or polynucleotide for the manufacture of a vaccine -  
XX Disclosure; Fig 1; 90pp; English.  
XX  
XX The present sequence represents a His-tagged ProT-D-Nef fusion protein.  
CC The protein is expressed in *Escherichia coli*, and is used to  
CC produce the vaccine of the invention. The specification describes  
CC the use of HIV Tat, HIV Nef, or Nef-Tat; and HIV gp120 in the  
CC manufacture of a vaccine. The vaccine is used for the prophylactic or  
CC therapeutic immunization of humans against HIV. Tat, Nef or Nef-Tat act  
CC in synergy with gp120 in the treatment and prevention of HIV. The  
CC vaccine reduces the HIV viral load in HIV infected humans and results  
CC in a maintenance of CD4+ levels over those levels found in the absence  
CC of vaccination with HIV Tat, Nef or Nef-Tat and HIV gp120.  
XX  
XX Sequence 326 AA;  
SQ

Query Match 67.5%; Score 1139.5; DB 22; Length 326;  
Best Local Similarity 71.2%; Pred. No. 2.2e-102;  
Matches 215; Conservative 0; Mismatches 0; Indels 87; Gaps 1;  
QY 1 MGKWKSSVVGWPTVRRMRRAEPAADGVGAASRDLEKHGATSSNTAATNAACAWLEA 60  
DB 112 mgkwwkssvvgwptvrrmrtaepaadvgaasrdlekhgaitssntaataacawlea 171  
QY 61 QEEEEVGFPTVPQVPLRPMTYKAAVDLSHFLKEKGGLEGLIHSQRQDILDLYHTQGY 120  
DB 172 qeeeevfgfptvpqvpplrmtykaavdlshflkekggleglshsqrqdlldlyhtqgy 231  
QY 121 FPDQWNTPGVGYRPLTFTGWCYKLVPEPDKVEEANKGENTSLHHPVSLHGMDPEREV 180  
DB 232 fpdwqntpgvgryplftfgwcyklvppepdkveeankgentslhhpvslhgmdperev 291  
QY 181 LEWRFSRLAFHHVARELHPEYFKNC----- 317  
DB 292 lewrfsrlafhhvarelhpeyfknc-----tsghhhh 324  
QY 241 CQVCFITKALGISYGRKKRRQRRPPQSGSOTHQVLSLKQPTSGRDPGPKETSGHHH 300  
DB 318 -----  
QY 301 HH 302  
DB 325 hh 326

RESULT 15  
AAY50795  
ID AAY50795 standard; protein; 206 AA.  
AC AAY50795;  
XX  
XX 17-FEB-2000 (first entry)  
DT Human NEF protein/calmodulin binding inhibitor.  
DE  
DE NEF protein/calmodulin; negative factor protein; binding inhibitor;  
KW diagnosis; detection; infection; treatment; HIV.  
KW  
XX Homo sapiens.  
XX  
XX WO957136-A2.  
PN  
XX 11-NOV-1999.  
PD  
XX 06-MAY-1999; 99WO-BP03105.  
XX  
XX 06-MAY-1998; 98DE-1020224.  
XX  
XX (SCHO/) SCHOTT M.  
PA

PA (SCHO/) SCHORR J.  
PA (ANTZ/) ANTZ C.  
XX  
XX Schott M, Schorr J, Antz C;  
XX WPI: 2000-038789/03.  
XX  
XX Binding agents used for treatment, prevention and diagnosis of human  
PT immune deficiency virus infection -  
PT  
XX Claim 3; Fig 1; 29pp; German.  
XX  
XX This invention describes novel binding partners (A) (i) for negative  
CC factor protein (Nef) that competitively inhibit binding of calmodulin (I)  
CC to Nef, or (ii) for (I) that competitively inhibit binding of Nef to (i).  
CC (A) are used for prevention, diagnosis (by specific detection of intra-  
CC and/or extra-cellular Nef, including staging of infection), and/or  
CC treatment of infections with human immune deficiency virus (HIV).  
XX  
XX Sequence 206 AA;  
SQ

Query Match 66.4%; Score 1120; DB 21; Length 206;  
Best Local Similarity 99.0%; Pred. No. 9.3e-101;  
Matches 204; Conservative 0; Mismatches 2; Indels 0; Gaps 0;  
QY 1 MGKWKSSVVGWPTVRRMRRAEPAADGVGAASRDLEKHGATSSNTAATNAACAWLEA 60  
DB 1 mgkwwkssvvgwptvrrmrtaepaadvgaasrdlekhgaitssntaannaacawlea 60  
QY 61 QEEEEVGFPTVPQVPLRPMTYKAAVDLSHFLKEKGGLEGLIHSQRQDILDLYHTQGY 120  
DB 61 qeeeevfgfptvpqvpplrmtykaavdlshflkekggleglshsqrqdlldlyhtqgy 120  
QY 121 FPDQWNTPGVGYRPLTFTGWCYKLVPEPDKVEEANKGENTSLHHPVSLHGMDPEREV 180  
DB 121 fpdwqntpgvgryplftfgwcyklvppepdkveeankgentslhhpvslhgmdperev 180  
QY 181 LEWRFSRLAFHHVARELHPEYFKNC 206  
DB 181 lewrfsrlafhhvarelhpeyfknc 206

RESULT 16  
AAR38893  
ID AAR38893 standard; Protein; 206 AA.  
XX  
XX AAR38893;  
XX  
XX 10-NOV-1993 (first entry)  
DT Nef protein of HIV-1.  
DE  
DE AIDS; antibody; p25; gp110; gp41; assay; detection;  
KW immunity; vaccine.  
KW  
XX Human immunodeficiency virus-1.  
XX  
XX US5221610-A.  
PN  
XX 22-JUN-1993.  
PD  
XX 26-MAY-1988; 88US-0199143.  
XX  
XX 26-MAY-1988; 88US-0199143.  
PR  
XX 04-SEP-1991; 91US-0754300.  
XX  
XX (INRM ) INST NAT SANTE & RECH MEDICALE.  
XX (INSP ) INST PASTEUR.  
XX  
XX Bahraoui EM, Chamaret S, Ferris S, Granier C, Montagnier L;  
PI Rietschoten JV, Rochat H, Sabatier JM;  
XX

DR WPI; 1993-213434/26.  
 XX  
 XX  
 PT Diagnosis of HIV infection - by detecting HIV antibodies using  
 PT antigenic polypeptide derived from nef protein of HIV-1  
 XX  
 XX  
 PS Disclosure; Fig 2; 15pp; English.  
 XX  
 XX  
 CC The nef protein comprises peptides which are expressed in vivo in HIV  
 CC infected patients before detectable amts. of p25, gp110 and gp41 are  
 CC expressed. Thus, they can be used in assays for early detection of HIV.  
 CC They can also be used to raise antibodies for use in detection,  
 CC to induce cellular immunity or to raise neutralising antibodies  
 CC that either inactivate the AIDS virus or reduce the viability of  
 CC the virus in vivo or destroy infected cells.  
 CC The peptides may be used in viral vaccines.  
 XX  
 XX  
 SQ Sequence 206 AA;

Query Match 66.1%; Score 1116; DB 14; Length 206;  
 Best Local Similarity 99.0%; Pred. No. 2.3e-100;  
 Matches 204; Conservative 0; Mismatches 2; Indels 0; Gaps 0;  
 QY 1 MGKWSKSSVGVGPTVREMRRAEPAADGVGAASRDLEKHGALTSSNTAATNAACAWLEA 60  
 DB 1 mggkwsksvvgvptvremrrraepaadvgaasrdlekhgaltssntaataacawlea 60  
 QY 61 QEEEEVGFPTPOVPLRPMTYKAADVLSHFLKEKGLGEGLIHSQRQDILDLMYHTQGY 120  
 DB 61 qeeeevfgptpovplrpmtykaadvlsflkkgleglshqrrqdildlwyhtqgy 120  
 QY 121 FPDWQNTTGPVGRYPLTFGCWCYKLYPVEPKVVEEANKGENTSLLHPVSLHGMDPPEREV 180  
 DB 121 fpdwqnttgpvgryppltfgwcylkypvepkvveeankgentsllhpvsllhgmdpper 180  
 QY 181 LEWRFDSRLAFHHVARELHPEYFKNC 206  
 DB 181 lewrfdsrlafhhvarelhpeyfknc 206

RESULT 17  
 AAB10054  
 ID AAB10054 standard; Protein; 206 AA.  
 XX  
 AC AAB10054;  
 XX  
 DT 02-NOV-2000 (first entry)  
 XX  
 DE HIV-1 nef protein.  
 XX  
 KW Glycoprotein; gag; pol; Gp-1; Gp-2; anti-HIV; cytostatic;  
 KW gene therapy; treatment; infectious disease; HIV; AIDS; neoplasm;  
 KW Carcinoma; melanoma; nef protein.  
 XX  
 OS Human immunodeficiency virus type 1.  
 XX  
 PN EP1006196-A2.  
 XX  
 PD 07-JUN-2000.  
 XX  
 PF 25-NOV-1999; 99EP-0250415.  
 XX  
 PR 26-NOV-1998; 98DE-1056463.  
 XX  
 XX (PETT-) PETTE INST HEINRICH.  
 XX  
 XX Von Laer MD;  
 PI  
 XX WPI; 2000-378268/33.  
 DR  
 DR N-PSDB; AAA40298, AAB10053, AAB10054.  
 XX  
 XX  
 PT New retroviral packing cell useful as pharmaceutical carrier in gene  
 PT therapy for treatment of HIV and neoplasms, comprises retroviral genes

PT and glycoproteins -  
 XX  
 PS Disclosure; Page 48; 69pp; German.  
 XX  
 CC This invention describes a novel retroviral packing cell (I), comprising  
 CC the retroviral genes gag, pol and glycoproteins GP-1 and GP-2 of the LCMV  
 CC coding gene gp, or a part of these. The products of the invention have  
 CC anti-HIV and cytostatic activity and can be used for gene therapy. (I) is  
 CC useful for in vitro infection of cells, especially hematopoietic stem  
 CC cells, for expression of transgenes in cells and as a pharmaceutical  
 CC carrier for gene therapy. (I) is therefore useful in the treatment of  
 CC infectious diseases (HIV or AIDS), neoplasms (carcinomas), melanomas and  
 CC other diseases. This sequence represents the Human immunodeficiency virus  
 CC (HIV-1) nef protein described in the method of the invention.  
 XX  
 XX  
 SQ Sequence 206 AA;

Query Match 66.1%; Score 1115; DB 21; Length 206;  
 Best Local Similarity 98.1%; Pred. No. 2.8e-100;  
 Matches 202; Conservative 1; Mismatches 3; Indels 0; Gaps 0;  
 QY 1 MGKWSKSSVGVGPTVREMRRAEPAADGVGAASRDLEKHGALTSSNTAATNAACAWLEA 60  
 DB 1 mggkwsksvvgvptvremrrraepaadvgaasrdlekhgaltssntaanaacawlea 60  
 QY 61 QEEEEVGFPTPOVPLRPMTYKAADVLSHFLKEKGLGEGLIHSQRQDILDLMYHTQGY 120  
 DB 61 qeeeevfgptpovplrpmtykaadvlsflkkgleglshqrrqdildlwyhtqgy 120  
 QY 121 FPDWQNTTGPVGRYPLTFGCWCYKLYPVEPKVVEEANKGENTSLLHPVSLHGMDPPEREV 180  
 DB 121 fpdwqnttgpvgryppltfgwcylkypvepkvveeankgentsllhpvsllhgmdpper 180  
 QY 181 LEWRFDSRLAFHHVARELHPEYFKNC 206  
 DB 181 lewrfdsrlafhhvarelhpeyfknc 206

RESULT 18  
 AAW89326  
 ID AAW89326 standard; Protein; 206 AA.  
 XX  
 AC AAW89326;  
 XX  
 DT 01-JUN-1999 (first entry)  
 XX  
 DE HIV-1 nef protein sequence.  
 XX  
 KW Antigenic composition; primate; lentivirus; nef gene; vaccine;  
 KW infection; AIDS; HIV-1; nef protein.  
 XX  
 OS Human immunodeficiency virus type 1.  
 XX  
 PN US5851813-A.  
 XX  
 PD 22-DEC-1998.  
 XX  
 PF 27-JAN-1994; 94US-0188583.  
 XX  
 PR 27-JAN-1994; 94US-0188583.  
 PR 12-JUL-1990; 90US-0551945.  
 PR 09-JUL-1991; 91US-0727494.  
 XX  
 XX (HARD ) HARVARD COLLEGE.  
 XX  
 XX Desrosiers RC;  
 PI  
 XX WPI; 1999-080408/07.  
 DR  
 DR N-PSDB; AAV81866.  
 XX  
 XX  
 PT Lentivirus antigenic compositions - containing lentivirus with nef  
 PT gene deletion

XX Disclosure; Fig 2A-R; 93pp; English.

XX The invention relates to an antigenic composition comprising an isolated

XX primate lentivirus whose genome contains an engineered non-revertible

XX null mutation in the nef gene, or an infectious DNA clone in a carrier.

XX The antigenic composition is used in vaccines against infection by the

XX lentivirus, e.g. AIDS.

SQ Sequence 206 AA;

Query Match 66.0%; Score 1114; DB 20; Length 206;

Best Local Similarity 97.6%; Pred. No. 3.5e-100;

Matches 201; Conservative 3; Mismatches 2; Indels 0; Gaps 0;

QY 1 MGGKWSKSSVGVPTVRMRRAEPAADGVGAASRDLEKHGAISSNTAATNAACAWLEA 60

DB 1 MGGKWSKSSVGVPTVRMRRAEPAADGVGAASRDLEKHGAISSNTAATNAACAWLEA 60

QY 61 QEEEEVGFPVTPQVPLRPMTYKAAVDLSHFLKEKGLGSLIHSORRODILDWIYHTQGY 120

DB 61 QEEKEVGFPVTPQVPLRPMTYKAAVDLSHFLKEKGLGSLIHSORRODILDWIYHTQGY 120

QY 121 FPDWQNTYTPGVRYPLTFGCWYKLVPEPKVEEANKGENTSLHPVSLHGMDDPEREV 180

DB 121 FPDWQNTYTPGVRYPLTFGCWYKLVPEPKVEEANKGENTSLHPVSLHGMDDPEREV 180

QY 181 LEWRDRLAFHVAHRELHPEYFKNC 206

DB 181 LEWRDRLAFHVAHRELHPEYFKNC 206

RESULT 19

AAP61515

ID AAP61515 standard; Protein; 206 AA.

XX AAP61515;

XX 08-JUN-1991 (first entry)

DE Sequence of E' protein.

XX HIV; LAV; AIDS; diagnosis; vaccine.

XX HTLV-IIIB/H9 cells (ATCC CRL 8543).

XX EP187041-A.

XX 09-JUL-1986.

XX 23-DEC-1985; 85EP-0309454.

XX 24-DEC-1984; 84US-0685272.

XX 04-DEC-1985; 85US-0805069.

XX (GETH ) GENENTECH INC.

XX Capon DJ, Lasky LA;

XX WPI; 1986-177602/28.

XX N-PSDB; AAN60288.

XX Acquired immune deficiency syndrome polypeptide(s) - obtd. by

XX molecular cloning etc. and used for diagnosis and in vaccines

XX against virus disease

XX Example; fig 2; 125pp; English.

XX A comparison of AAN60287 with the cDNA of the HTLV-III genome

XX revealed one particular clone, designated p7.11 which contained a

XX DNA sequence encoding this peptide (AAP60308) sequence. This approx.

XX 2.2 kilobase covers the precursor gag region and encodes, 5' to 3',

CC p-12, p-15, p-24 a second p-15 protein, and approx. 300 extra base

CC pairs 3' to the gag region (see AAN60288).

XX Sequence 206 AA;

Query Match 65.8%; Score 1111; DB 7; Length 206;

Best Local Similarity 97.1%; Pred. No. 7e-100;

Matches 200; Conservative 4; Mismatches 2; Indels 0; Gaps 0;

QY 1 MGGKWSKSSVGVPTVRMRRAEPAADGVGAASRDLEKHGAISSNTAATNAACAWLEA 60

DB 1 MGGKWSKSSVGVPTVRMRRAEPAADGVGAASRDLEKHGAISSNTAATNAACAWLEA 60

QY 61 QEEEEVGFPVTPQVPLRPMTYKAAVDLSHFLKEKGLGSLIHSORRODILDWIYHTQGY 120

DB 61 QEEKEVGFPVTPQVPLRPMTYKAAVDLSHFLKEKGLGSLIHSORRODILDWIYHTQGY 120

QY 121 FPDWQNTYTPGVRYPLTFGCWYKLVPEPKVEEANKGENTSLHPVSLHGMDDPEREV 180

DB 121 FPDWQNTYTPGVRYPLTFGCWYKLVPEPKVEEANKGENTSLHPVSLHGMDDPEREV 180

QY 181 LEWRDRLAFHVAHRELHPEYFKNC 206

DB 181 LEWRDRLAFHVAHRELHPEYFKNC 206

RESULT 20

AAP60423

ID AAP60423 standard; Protein; 216 AA.

XX AAP60423;

XX 20-AUG-1991 (first entry)

DE Sequence of LAV virus ORF F protein.

XX AIDS vaccine; diagnosis; immunoassay; HIV; HTLV-III.

XX Lymphadenopathy virus.

XX WO8602383-A.

XX 24-APR-1986.

XX 18-OCT-1985; 85WO-EP00548.

XX 21-JAN-1985; 85GB-0001473.

XX 18-OCT-1984; 84FR-0016013.

XX 16-NOV-1984; 84GB-0029099.

XX (CNRS ) CNRS CENT NAT RECH SCI.

XX (INSP ) INST PASTEUR.

XX Montagnier L, Krust B, Chamaret S, Clavel F, Chermann J-C;

XX Barre-Sinoussi F, Alizon M, Sonigo P, Stewart C, Danos O;

XX Wain-Hobson S;

XX WPI; 1986-119166/18.

XX N-PSDB; AAN60365.

XX Purified glyco:protein and peptide(s) - are recognised by sera contg.

XX antibodies against lymphadenopathy virus and useful in detecting

XX AIDS antibodies or in vaccines

XX Disclosure; Fig 4; 75pp; English.

XX The inventors claim a polypeptide which is recognised by sera of

XX human origin contg. antibodies against the virus of

XX lymphadenopathies (LAV) or acquired immune deficiency syndrome

XX (AIDS). Also claimed are various peptides corresp. to the AA

XX sequences deducible from proteins encoded by LAV DNA, defined by

XX specific residues (e.g. 12-32, 37-46, 49-79, 88-153) in accordance

CC with a formula given in the specification.

```
XX SQ Sequence 216 AA;
Query Match 65.68; Score 1107; DB 7; Length 216;
Best Local Similarity 98.18; Pred. No. 1.8e-99;
Matches 202; Conservative 1; Mismatches 3; Indels 0; Gaps 0;
Qy 1 MGGKWSKSVVGPVTRMRRAEPAADGVGAASRDLEKKGAISSNTAATNAACAWLEA 60
Db 11 mggkwsksvvgpvtvrmtraepaahgvaaspdlkkgaitssntaataacawlea 70
Qy 61 QEEVEGFPVTPQVPLRPMYKAAVDLSHFLKEKGGLEGLHSQRQDILDWIYHTGGY 120
Db 71 qeeevgfpvtpqvpplrmpykaaavdlshflkekggleglshsgrqrdildwihtggy 130
Qy 121 PPDQNYTPGCVRYPLTFGWCYKLVPEPKVEANKGENTSLHPVSLHGMDDPREV 180
Db 131 fpdwqnytpgpgvrypltfwgcyklvppepkveeankgentsllhpsvslhgmddperev 190
Qy 181 LEWRDRLAFHHVARELHPYFKNC 206
Db 191 lewrfdsrflafhhvarelhpqyfknc 216
RESULT 21
AAW90179
ID AAW90179 standard; Protein; 206 AA.
AC AAW90179;
XX
DT 14-JUN-1999 (first entry)
XX
DE HTLV-III E' protein.
XX
KW AIDS; acquired immunodeficiency syndrome; viral infection; env protein;
KW fusion polypeptide; E' protein; gag protein; pol protein; P' protein;
KW gp signal peptide; detection; vaccination; etiological agent;
KW infection inhibitor; AIDS associated retrovirus; p24; gp41.
XX
OS Human lymphotropic virus type III.
XX
FH Key Location/Qualifiers
FT Misc-difference 11 /label= Val, Ile
FT FT Misc-difference 51 /label= Thr, Asn
FT FT Misc-difference 54 /label= Asp, Ala
FT FT Misc-difference 65 /label= Glu, Lys
FT FT Misc-difference 151 /label= Glu, Asp
FT FT Misc-difference 153 /label= Leu, Val
XX
PN US5853978-A.
XX
PD 29-DEC-1998.
XX
PF 29-JUL-1994; 94US-0282857.
XX
PR 08-MAY-1986; 86US-0861016.
PR 24-DEC-1984; 84US-0685272.
PR 04-DEC-1985; 85US-0805069.
PR 02-AUG-1988; 88US-0227568.
PR 19-NOV-1992; 92US-0979391.
PR 29-SEP-1993; 93US-0129009.
PR 29-JUL-1994; 94US-0282857.
XX
PA (GETH ) GENENTECH INC.
XX
```

```
PI Berman PW, Capon DJ, Lasky LA;
XX WPI; 1999-094894/08.
DR N-PSDB; AAV74271.
XX
PT New AIDS-associated fusion polypeptide(s) - used for the detection
XX Of AIDS or to inhibit infection by AIDS associated retrovirus or
XX dissemination of such retrovirus in infected individuals
XX
PS Example 1; Fig 2A-J; 47pp; English.
XX
CC This invention describes a novel fusion polypeptide which comprises
CC (a) a first polypeptide sequence of an AIDS associated E', env, or gag
CC polypeptide, that specifically binds complementary antibody and (b) a
CC second polypeptide sequence which is not an AIDS associated virus
CC polypeptide. Also described are (1) a fusion polypeptide having a
CC herpes simplex virus (HSV) gp signal peptide sequence fused in reading
CC frame with a polypeptide sequence other than HSV gp, (2) a nucleic acid
CC encoding a fusion polypeptide as in (1); (3) an expression vector
CC comprising a nucleic acid as in (2); (4) a host comprising a vector as
CC in (3). The AIDS-associated fusion polypeptides can be used in the
CC detection of and vaccination against viral etiological agents of AIDS.
CC They can also be administered as a pharmaceutical agent to inhibit
CC infection by AIDS associated retrovirus or dissemination of such
CC retrovirus in infected individuals.
XX
SQ Sequence 206 AA;
Query Match 64.7%; Score 1092; DB 20; Length 206;
Best Local Similarity 96.18; Pred. No. 4.9e-98;
Matches 198; Conservative 1; Mismatches 7; Indels 0; Gaps 0;
Qy 1 MGGKWSKSVVGPVTRMRRAEPAADGVGAASRDLEKKGAISSNTAATNAACAWLEA 60
Db 1 mggkwsksvvgpvtvrmtraepaahgvaaspdlkkgaitssntaataacawlea 60
Qy 61 QEEVEGFPVTPQVPLRPMYKAAVDLSHFLKEKGGLEGLHSQRQDILDWIYHTGGY 120
Db 61 qeeevgfpvtpqvpplrmpykaaavdlshflkekggleglshsgrqrdildwihtggy 120
Qy 121 PPDQNYTPGCVRYPLTFGWCYKLVPEPKVEANKGENTSLHPVSLHGMDDPREV 180
Db 121 fpdwqnytpgpgvrypltfwgcyklvppepkveeankgentsllhpsvslhgmddperev 180
Qy 181 LEWRDRLAFHHVARELHPYFKNC 206
Db 181 lewrfdsrflafhhvarelhpqyfknc 206
RESULT 22
AAE04960
ID AAE04960 standard; Protein; 216 AA.
XX
AC AAE04960;
XX
DT 10-SEP-2001 (first entry)
XX
DE HIV-1 jrf1 Nef protein.
XX
KW DNA vaccine; codon optimisation; Nef protein; HIV infection; anti-HIV;
KW cell mediated immune response; cytotoxic T lymphocyte; CTL;
KW human immunodeficiency virus-1; HIV-1; mutant; muten.
XX
OS Human immunodeficiency virus type 1.
XX
SY Synthetic.
XX
FH Key Location/Qualifiers
FT Modified-site 2 /note= "Myristylation site"
XX
PN WO200143693-A2.
XX
```

PD 21-JUN-2001.  
 XX  
 XX 15-DEC-2000; 2000WO-US34162.  
 PF  
 XX 17-DEC-1999; 99US-0172442.  
 PR  
 XX (MERI ) MERCK & CO INC.  
 PA  
 XX Shiver JW, Liang X, Fu T;  
 PI  
 XX WPI; 2001-417878/44.  
 DR N-PSDB; AAD09603.  
 DR  
 XX Novel DNA vaccine useful for inducing cell mediated immune response  
 PT against virulent strains of HIV, comprises DNA expression vector, and  
 PT DNA containing codon optimized open reading frame encoding Nef protein  
 PT  
 PT  
 XX  
 PS Claim 5; Page 13; 84pp; English.  
 PS  
 XX The present invention relates to a DNA vaccine comprising a DNA  
 CC expression vector, and a DNA molecule containing a codon optimised open  
 CC reading frame encoding a Nef protein or its immunogenic derivative. The  
 CC Nef protein or its derivative is expressed and generates an immune  
 CC response which provides a substantial level of protection against HIV  
 CC infection, upon administration of to a host. The DNA vaccine is useful  
 CC for inducing a cell mediated immune (cytotoxic T lymphocyte, CTL)  
 CC response against infection or disease caused by virulent strains of HIV,  
 CC by administering the DNA vaccine into the tissue of the vertebrate host,  
 CC preferably human. It is also useful for lowering transmission rate to  
 CC previously uninfected individuals and/or for reducing levels of viral  
 CC loads within an infected individual, so as to prolong the asymptomatic  
 CC phase of HIV infection. The present sequence is human immunodeficiency  
 CC virus-1 (HIV-1) Jrf1 Nef protein. The different codon optimised nef  
 CC constructs of the invention are used in the exemplification of the  
 CC invention.  
 XX  
 SQ Sequence 216 AA;  
  
 Query Match 59.9%; Score 1011; DB 22; Length 216;  
 Best Local Similarity 83.8%; Pred. No. 4e-90;  
 Matches 181; Conservative 15; Mismatches 10; Indels 10; Gaps 1;  
  
 QY 1 MGKSKSSVVGWPTVRRMRRAEPAAD-----GVGAASRDLEKHGAISSNTAA 50  
 DB 1 mggkswksrsvpgwstvrermrraeapaadrvttepaavgvgsrdlekghaltsntaa 60  
  
 QY 51 TNAACAWLEAQEEVEGFPVTPQVPLRPMTRYKAAVDLSHFLKEKGLGLHSORRQDIL 110  
 DB 61 tnadacawleaqeadeevgfvrpgvplrpmtkygavdlsflkkggleglihsqkqdl 120  
  
 QY 111 DLMIYHTQGYFPDQWNYTPGPGVRYPLTFGWCYKLVPEPDKVEEANKGENTSLLHPVSL 170  
 DB 121 dlwvyhtqgyfwdqnytpgpgirfltfgwckfklvpepekeveanegenncillhpsmq 180  
  
 QY 171 HGMDPDERVLEWRDLSRLAFHVAHRELHPEYKNC 206  
 DB 181 hgtedpekeviewrfosklafhvharelhpeyykdc 216  
  
 RESULT 23  
 AAW53113  
 ID AAW53113 standard; Protein; 210 AA.  
 XX  
 AC AAW53113;  
 XX  
 DT 25-JUN-1998 (first entry)  
 XX  
 DE Protein 6 contained in a complete ARV-2 nucleotide sequence.  
 XX  
 XX ARV-2; enhanced promoter; gene expression; cytomegalovirus;  
 KW HIV; AIDS.  
 XX

XX Human immunodeficiency virus type 1.  
 OS  
 XX US5688688-A.  
 PN  
 XX 18-NOV-1997.  
 PD  
 XX  
 PF 10-AUG-1994; 94US-0288336.  
 XX  
 PR 24-DEC-1987; 87US-0138894.  
 PR 31-OCT-1984; 84US-0667501.  
 PR 30-JAN-1985; 85US-0696534.  
 PR 06-SEP-1985; 85US-0773447.  
 PR 17-AUG-1992; 92US-0931191.  
 PR 28-JUN-1993; 93US-0083391.  
 PR 17-AUG-1993; 93US-0107377.  
 PR 10-AUG-1994; 94US-0288336.  
 XX  
 PA (CHIR ) CHIRON CORP.  
 XX  
 XX Chapman BS, Dina D, Haigwood NL, Luciw PA, Rosenberg S;  
 PI Thayer RM;  
 XX  
 XX WPI; 1998-007982/01.  
 XX  
 PT Enhanced promoter for gene expression - comprising cytomegalovirus  
 PT immediate early promoter plus intron  
 PT  
 PS Example 1; Fig 4C-P; 99pp; English.  
 XX  
 CC This sequence represents a protein of unspecified function contained in a  
 CC complete nucleotide sequence of ARV-2 derived from partial sequences of  
 CC several ARV clones. The invention provides a method for construction of a  
 CC vector for expression of a polypeptide in a mammalian cell, comprising a  
 CC polypeptide coding sequence operably linked downstream of an enhanced  
 CC promoter. The enhanced promoter comprises the human cytomegalovirus  
 CC immediate early region (HCMV IE1) promoter and the first intron proximate  
 CC to the 3' end of the HCMV IE1 promoter. The polypeptide can be any of the  
 CC HIV recombinant polypeptides and especially HIV gp120. Expression of HIV  
 CC gp120 by COS 7 cells transfected with pcMV6a containing the above enhanced promoter,  
 CC region, where pcMV6a is a vector containing the above enhanced promoter,  
 CC is increased by a factor of 50-100 compared with the use of a vector  
 CC containing the SV40 early promoter.  
 XX  
 SQ Sequence 210 AA;  
  
 Query Match 59.1%; Score 997; DB 19; Length 210;  
 Best Local Similarity 85.2%; Pred. No. 8.9e-89;  
 Matches 179; Conservative 16; Mismatches 11; Indels 4; Gaps 1;  
  
 QY 1 MGKSKSSVVGWPTVRRMR-----RAEPAADGVGAASRDLEKHGAISSNTAATNACA 56  
 DB 1 mggkswksrsmggwsalrermrraeapaadvgavsrdrlekghaltsntaatnadca 60  
  
 QY 57 WLEAQEEVEGFPVTPQVPLRPMTRYKAAVDLSHFLKEKGLGLHSORRQDILDLIYH 116  
 DB 61 wleaqeeveevgfvrpgvplrpmtkykaavdlsflkkggleglihsqrqellldliyh 120  
  
 QY 117 TQGYFPDQWNYTPGPGVRYPLTFGWCYKLVPEPDKVEEANKGENTSLLHPVSLHGMDP 176  
 DB 121 tqgyfwdqnytpgpgirfltfgwckfklvpepekeveanegenncillhpsmlhgmada 180  
  
 QY 177 EREVLEWRDLSRLAFHVAHRELHPEYKNC 206  
 DB 181 ekevlvwrfdsklafhvharelhpeyykdc 210  
  
 RESULT 24  
 AAY77299  
 ID AAY77299 standard; Protein; 210 AA.  
 XX  
 AC AAY77299;

XX DT 22-MAY-2000 (first entry)  
XX DE HIV-1 (ATCC CRL 8597) ORF #3-encoded protein.  
XX KW HIV-1; immunoassay; antigen; pol fragment; p31; recombinant; antibody;  
XX KW detection.  
XX OS Human immunodeficiency virus type 1 'ATCC CRL 8597'.  
XX PN US6013432-A.  
XX PD 11-JAN-2000.  
XX PF 17-MAY-1995; 95US-0443434.  
XX PR 08-JUL-1993; 93US-0089407.  
XX PR 24-DEC-1987; 87US-0138894.  
XX PR 17-AUG-1992; 92US-0931154.  
XX PR 31-OCT-1984; 84US-0667501.  
XX PR 30-JAN-1985; 85US-0696534.  
XX PR 06-SEP-1985; 85US-0773447.  
XX (CHIR ) CHIRON CORP.  
XX LUciw PA, Dina D;  
XX WPI; 2000-170256/15.  
XX N-PSDB; AA290201.  
XX Immunoassay for antibodies against human immune deficiency virus, for  
XX diagnosing infection, uses an immunogenic fragment of the pol protein  
XX as antigen -  
XX Example 1; Fig 40-P; 99pp; English.  
XX The invention relates to the improvement of HIV-1 immunoassays by the  
XX use of an HIV-1 antigen comprising an immunogenic fragment of  
XX recombinant or synthetic HIV-1 pol, which is encoded by an approximately  
XX 9.7 kb sequence between a BstXI restriction site at position 3006 and an  
XX NdeI site at position 5131 of the genome (the proviral DNA sequence is  
XX given in AA290201). The immunogenic pol fragment is not immunologically  
XX cross-reactive with human T cell lymphotropic viruses I or II. The  
XX invention also encompasses the use of p31 as an antigen. The recombinant  
XX antigens may be produced in Escherichia coli, Saccharomyces cerevisiae or  
XX in mammalian cells. Immunoassays using the recombinant HIV proteins may  
XX be used to diagnose and stage HIV-1 infections. Sequences  
XX AA27294-Y77299 represent proteins encoded by the genome of HIV-1 (ATCC  
XX CRL 8597).  
XX Sequence 210 AA;  
XX SQ  
Query Match 59.1%; Score 997; DB 21; Length 210;  
Best Local Similarity 85.2%; Pred. No. 8.9e-89;  
Matches 179; Conservative 16; Mismatches 11; Indels 4; Gaps 1;  
QY 1 MGGKWSKSSVGVGPTVRMRMR-----RAEPAADGVGAASRDLEKHGAITSNTAATNAACA 56  
Db 1 mggkwsksmggwsaiaermtaepaepaagdgavgsrdlekhgaitssntaataadca 60  
QY 57 WLEAQEEVEGFVPTQVPLRPMYKAAVDLSHFLKEKGLLEGLIHSQRQDILDWIVH 116  
Db 61 wleaqeevegfvrpqplrptmykaalishflkekgglegllwsqrqeldlwiyh 120  
QY 117 TQGYFDDWNYTPGCVRYPLTFGCYKLVPEPDKVEEANKGENTSLHHPVSLHGMDDP 176  
Db 121 tqgyfddwnytpgcvrypltfgcycylvpepekdveeankgenntslhphmslhgmeda 180  
QY 177 EREVLEWRFDSRLAFHVAHRELHPEYFKNC 206  
Db 181 ekevlwrfdsrlafhvharelhpeyfkdc 210

RESULT 25  
AAE04962  
ID AAE04962 standard; Protein; 217 AA.  
XX AC AAE04962;  
XX DT 10-SEP-2001 (first entry)  
XX DE HIV-1 jrf1 Nef (G2A, LLAA) mutant protein.  
XX KW DNA vaccine; codon optimisation; Nef protein; HIV infection; anti-HIV;  
XX KW cell mediated immune response; cytotoxic T lymphocyte; CTL; mutein.  
XX KW human immunodeficiency virus-1; HIV-1; jrf1 Nef; mutant; mutein.  
XX OS Human immunodeficiency virus type 1.  
XX Synthetic.  
XX Key Location/Qualifiers  
XX FH Misc-difference 2 /note= "Wild type myristylation site at Gly2 substituted  
XX FT with Ala"  
XX FT Misc-difference 174..175 /note= "Wild type dileucine motif (Leu-Leu) substituted  
XX FT with Ala-Ala"  
XX PN WO200143693-A2.  
XX PD 21-JUN-2001.  
XX PF 15-DEC-2000; 2000WO-US34162.  
XX PR 17-DEC-1999; 99US-0172442.  
XX PA (MERI ) MERCK & CO INC.  
XX PI Shiver JW, Liang X, Fu T;  
XX WPI; 2001-417878/44.  
XX N-PSDB; AAD09605.  
XX Novel DNA vaccine useful for inducing cell mediated immune response  
XX against virulent strains of HIV, comprises DNA expression vector, and  
XX DNA containing codon optimized open reading frame encoding Nef protein  
XX Claim 29; Page 18; 84pp; English.  
XX The present invention relates to a DNA vaccine comprising a DNA  
XX expression vector, and a DNA molecule containing a codon optimised open  
XX reading frame encoding a Nef protein or its immunogenic derivative. The  
XX Nef protein or its derivative is expressed and generates an immune  
XX response which provides a substantial level of protection against HIV  
XX infection, upon administration of to a host. The DNA vaccine is useful  
XX for inducing a cell mediated immune (cytotoxic T lymphocyte, CTL)  
XX response against infection or disease caused by virulent strains of HIV,  
XX by administering the DNA vaccine into the tissue of the vertebrate host,  
XX preferably human. It is also useful for lowering transmission rate to  
XX previously uninfected individuals and/or for reducing levels of viral  
XX loads within an infected individual, so as to prolong the asymptomatic  
XX phase of HIV infection. The present sequence is human immunodeficiency  
XX virus-1 (HIV-1) jrf1 Nef (G2A, LLAA) mutant protein. The different nef  
XX constructs of the invention are used in the exemplification of the  
XX SQ Sequence 217 AA;  
Query Match 59.0%; Score 996; DB 22; Length 217;  
Best Local Similarity 82.0%; Pred. No. 1.2e-88;  
Matches 178; Conservative 16; Mismatches 13; Indels 10; Gaps 1;  
QY 1 MGGKWSKSSVGVGPTVRMRMRRAEPAAD-----GVGAASRDLEKHGAITSNTAA 50



1 magkkskrsvpgwstvrermttaepaadvrrtpeaavgvgsrdlekghaitssntaa 60  
51 TNAACAWLEAQEEVEGFVPTQVPLRPMTYKAAVDLSHFLKEKGLGSLHSQRQDIL 110  
61 tnadcawleaqeadeevgfvprpqvplrpmtkygavdlshflkekglgslhsqkrqdl 120  
111 DLWIVHTQGYFDPWQNTYTPGVRYPVLTGWCYKLVPEPDKVEEANKGENTSLHHPVSL 170  
121 dlwvhtqgyfdpwqntypggyrlrftlgwcfklvppekeveeaneenncaahpmsq 180  
171 HGMDPPEVLEWRFDSRLAFHHVARELHPEYFKNCT 207  
181 hgiedpekevelewfdsrlafhhvarelhpeyykdc 217

RESULT 26  
AA12262  
ID AAR12262 standard; Protein; 210 AA.  
AC AAR12262;  
XX 20-AUG-1991 (first entry)  
XX  
DE HIV-1 strain OYI open reading frame (ORF) F protein.  
XX  
KW HIV-1; AIDS; retroviruses.  
XX  
OS Homo sapiens.  
XX  
XX US5019510-A.  
XX 28-MAY-1991.  
XX  
XX 28-OCT-1987; 87US-0113655.  
XX  
XX 28-OCT-1987; 87US-0113655.  
XX  
XX (INSP ) INST PASTEUR.  
XX  
XX Wain-Hobson S, Huet T, Delaporte E, Brun-Vezinet F;  
XX WPI; 1991-177518/24.  
XX  
XX Purified human retrovirus - is mutant of HIV-1 having  
XX characteristics of HIV-1 OYI, used in diagnosis of HIV infection  
XX  
XX Disclosure; fig 8; 23pp; English.  
XX  
XX This sequence constitutes the ORF F protein constituent of a new  
XX strain of HIV-1 retrovirus, OYI. This mutant retroviral strain is  
XX useful in an assay for diagnosing HIV infection. See also AAR11943  
XX (OYI nucleotide sequence), AAR12255-61 (other HIV OYI constituent  
XX proteins).  
XX  
XX Sequence 210 AA;

Query Match 58.2%; Score 983; DB 12; Length 210;  
Best Local Similarity 82.5%; Pred. NO. 2e-87;  
Matches 174; Conservative 22; Mismatches 9; Indels 6; Gaps 2;  
QY 1 MCGKWSKSSVVGHTVREMRRA-----EPAADGVGAASRDLEKKGAGTSSNTAATNAAC 55  
DB 1 mggkwsksmkgvptiremrkralpppeaagvgvaasrdlekghaitssntaatnadc 60  
QY 56 AWLEAQEEVEGFVPTQVPLRPMTYKAAVDLSHFLKEKGLGSLHSQRQDILDWIY 115  
DB 61 awleaqeadeevgfvprpqvplrpmtkygavdlshflkekglgslhsqkrqdlld-wvy 119  
QY 116 HTGQYFDPWQNTYTPGVRYPVLTGWCYKLVPEPDKVEEANKGENTSLHHPVSLHGMDD 175  
DB 120 htggyfdpwqntypggyrlrftlgwcfklvpmdpdqveeaneennslhlpishgmdd 179

QY 176 PEREVLWRFDSRLAFHHVARELHPEYFKNC 206  
DB 180 pekevlwrfdsrlafhhvarelhpeyykdc 210  
RESULT 27  
AAE04961  
ID AAE04961 standard; Protein; 237 AA.  
XX  
AC AAE04961;  
XX  
DT 10-SEP-2001 (first entry)  
XX  
DE Human tPA leader peptide-HIV-1 jrf1 Nef (opt tpanef) mutant protein.  
XX  
KW DNA vaccine; codon optimisation; Nef protein; HIV infection; anti-HIV;  
KW cell mediated immune response; cytotoxic T lymphocyte; CTL;  
KW human immunodeficiency virus-1; HIV-1; human; jrf1 Nef; opt tpanef;  
KW tissue plasminogen activator; tPA; chimeric protein; mutant; mutein.  
XX  
OS Chimeric - Human immunodeficiency virus type 1.  
OS Chimeric - Homo sapiens.  
XX  
XX  
FH Key Location/Qualifiers  
FT Peptide 1..26  
FT /label= Leader\_peptide  
FT /note= "Human tissue plasminogen activator (tPA)  
FT leader sequence"  
FT Protein 27..237  
FT /note= "Mature HIV-1 Nef (6-216aa) protein"  
FT Domain 195..196  
FT /label= Dileucine\_motif  
XX  
XX WO200143693-A2.  
XX  
XX 21-JUN-2001.  
XX  
XX 15-DEC-2000; 2000WO-US34162.  
XX  
XX 17-DEC-1999; 99US-0172442.  
XX  
XX (MERI ) MERCK & CO INC.  
XX  
XX Shiver JW, Liang X, Fu T;  
XX  
XX WPI; 2001-417878/44.  
XX  
XX N-PSDB; AAD09604.  
XX  
XX  
XX Novel DNA vaccine useful for inducing cell mediated immune response  
XX against virulent strains of HIV, comprises DNA expression vector, and  
XX DNA containing codon optimized open reading frame encoding Nef protein  
XX  
XX Claim 29; Page 17; 84pp; English.  
XX  
XX The present invention relates to a DNA vaccine comprising a DNA  
XX expression vector, and a DNA molecule containing a codon optimised open  
XX reading frame encoding a Nef protein or its immunogenic derivative. The  
XX Nef protein or its derivative is expressed and generates an immune  
XX response which provides a substantial level of protection against HIV  
XX infection, upon administration of to a host. The DNA vaccine is useful  
XX for inducing a cell mediated immune (cytotoxic T lymphocyte, CTL)  
XX response against infection or disease caused by virulent strains of HIV,  
XX by administering the DNA vaccine into the tissue of the vertebrate host,  
XX preferably human. It is also useful for lowering transmission rate to  
XX previously uninfected individuals and/or for reducing levels of viral  
XX loads within an infected individual, so as to prolong the asymptomatic  
XX phase of HIV infection. The present sequence is human tissue plasminogen  
XX activator (tPA) leader peptide-human immunodeficiency virus-1 (HIV-1)  
XX jrf1 Nef (opt tpanef) chimeric mutant protein. The mutation include  
XX deletion of the myristylation site (Gly2) in the HIV-1 jrf1 Nef protein.  
XX The different codon optimised nef constructs of the invention are used

CC in the exemplification of the invention.

XX Sequence 237 AA;

Query Match 57.9%; Score 978; DB 22; Length 237;

Best Local Similarity 83.4%; Pred. No. 7.4e-87;

Matches 176; Conservative 15; Mismatches 10; Indels 10; Gaps 1;

QY 6 SKSSVVGWPTVRMRRAEPAAD-----GVGAASRDLEKHAITSSNTAATNAAC 55

DB 27 skrspgstvtrmraraeaaadvrrtapaavgavsdldlekhaitsntaatnadc 86

QY 56 AWLEAQEEVEGFPVTPQVPLRPMTYKAAVDLSHFLKKGGLGSLHSORRQDILDIWY 115

DB 87 awleaqedeegfpvtrpqrplrpmtkygavdlshflikeggleglihsqrqildlwy 146

QY 116 HTQGYFPDQWNTYTPGCVRYPLTFGWCYKLVPEPKVEEANKGENTSLHPVSLHGMD 175

DB 147 htqgyfpdwqnytpggyirfpltfwgcfklvpvepkveeangenncllhpmshgied 206

QY 176 PEREVLWRFDLSRLAFHHVARELHPYFKNC 206

DB 207 pekeviewrfdsklathhvarelhpeyykdc 237

QY 176 PEREVLWRFDLSRLAFHHVARELHPYFKNC 206

DB 207 pekeviewrfdsklathhvarelhpeyykdc 237

QY 176 PEREVLWRFDLSRLAFHHVARELHPYFKNC 206

DB 207 pekeviewrfdsklathhvarelhpeyykdc 237

QY 176 PEREVLWRFDLSRLAFHHVARELHPYFKNC 206

DB 207 pekeviewrfdsklathhvarelhpeyykdc 237

QY 176 PEREVLWRFDLSRLAFHHVARELHPYFKNC 206

DB 207 pekeviewrfdsklathhvarelhpeyykdc 237

QY 176 PEREVLWRFDLSRLAFHHVARELHPYFKNC 206

DB 207 pekeviewrfdsklathhvarelhpeyykdc 237

QY 176 PEREVLWRFDLSRLAFHHVARELHPYFKNC 206

DB 207 pekeviewrfdsklathhvarelhpeyykdc 237

QY 176 PEREVLWRFDLSRLAFHHVARELHPYFKNC 206

DB 207 pekeviewrfdsklathhvarelhpeyykdc 237

QY 176 PEREVLWRFDLSRLAFHHVARELHPYFKNC 206

DB 207 pekeviewrfdsklathhvarelhpeyykdc 237

QY 176 PEREVLWRFDLSRLAFHHVARELHPYFKNC 206

DB 207 pekeviewrfdsklathhvarelhpeyykdc 237

QY 176 PEREVLWRFDLSRLAFHHVARELHPYFKNC 206

DB 207 pekeviewrfdsklathhvarelhpeyykdc 237

QY 176 PEREVLWRFDLSRLAFHHVARELHPYFKNC 206

DB 207 pekeviewrfdsklathhvarelhpeyykdc 237

QY 176 PEREVLWRFDLSRLAFHHVARELHPYFKNC 206

DB 207 pekeviewrfdsklathhvarelhpeyykdc 237

QY 176 PEREVLWRFDLSRLAFHHVARELHPYFKNC 206

DB 207 pekeviewrfdsklathhvarelhpeyykdc 237

QY 176 PEREVLWRFDLSRLAFHHVARELHPYFKNC 206

DB 207 pekeviewrfdsklathhvarelhpeyykdc 237

QY 176 PEREVLWRFDLSRLAFHHVARELHPYFKNC 206

DB 207 pekeviewrfdsklathhvarelhpeyykdc 237

QY 176 PEREVLWRFDLSRLAFHHVARELHPYFKNC 206

DB 207 pekeviewrfdsklathhvarelhpeyykdc 237

QY 176 PEREVLWRFDLSRLAFHHVARELHPYFKNC 206

DB 207 pekeviewrfdsklathhvarelhpeyykdc 237

QY 176 PEREVLWRFDLSRLAFHHVARELHPYFKNC 206

DB 207 pekeviewrfdsklathhvarelhpeyykdc 237

QY 176 PEREVLWRFDLSRLAFHHVARELHPYFKNC 206

DB 207 pekeviewrfdsklathhvarelhpeyykdc 237

QY 176 PEREVLWRFDLSRLAFHHVARELHPYFKNC 206

DB 207 pekeviewrfdsklathhvarelhpeyykdc 237

QY 176 PEREVLWRFDLSRLAFHHVARELHPYFKNC 206

DB 207 pekeviewrfdsklathhvarelhpeyykdc 237

QY 176 PEREVLWRFDLSRLAFHHVARELHPYFKNC 206

DB 207 pekeviewrfdsklathhvarelhpeyykdc 237

QY 176 PEREVLWRFDLSRLAFHHVARELHPYFKNC 206

DB 207 pekeviewrfdsklathhvarelhpeyykdc 237

QY 176 PEREVLWRFDLSRLAFHHVARELHPYFKNC 206

PT

XX

PS

XX

CC

CC

CC

CC

CC

CC

CC

CC

CC

CC

CC

CC

CC

CC

CC

CC

CC

CC

CC

CC

CC

CC

CC

CC

CC

CC

CC

CC

CC

CC

CC

CC

CC

CC

CC

CC

CC

CC

CC

CC

CC

CC

CC

CC

CC

CC

CC

CC

CC

CC

CC

CC

CC

CC

CC

CC

CC

CC

Query Match 57.3%; Score 968; DB 22; Length 237;

Best Local Similarity 82.5%; Pred. No. 7e-86;

Matches 174; Conservative 15; Mismatches 12; Indels 10; Gaps 1;

QY 6 SKSSVVGWPTVRMRRAEPAAD-----GVGAASRDLEKHAITSSNTAATNAAC 55

DB 27 skrspgstvtrmraraeaaadvrrtapaavgavsdldlekhaitsntaatnadc 86

QY 56 AWLEAQEEVEGFPVTPQVPLRPMTYKAAVDLSHFLKKGGLGSLHSORRQDILDIWY 115

DB 87 awleaqedeegfpvtrpqrplrpmtkygavdlshflikeggleglihsqrqildlwy 146

QY 116 HTQGYFPDQWNTYTPGCVRYPLTFGWCYKLVPEPKVEEANKGENTSLHPVSLHGMD 175

DB 147 htqgyfpdwqnytpggyirfpltfwgcfklvpvepkveeangenncllhpmshgied 206

QY 176 PEREVLWRFDLSRLAFHHVARELHPYFKNC 206

DB 207 pekeviewrfdsklathhvarelhpeyykdc 237

QY 176 PEREVLWRFDLSRLAFHHVARELHPYFKNC 206

DB 207 pekeviewrfdsklathhvarelhpeyykdc 237

QY 176 PEREVLWRFDLSRLAFHHVARELHPYFKNC 206

DB 207 pekeviewrfdsklathhvarelhpeyykdc 237

QY 176 PEREVLWRFDLSRLAFHHVARELHPYFKNC 206

DB 207 pekeviewrfdsklathhvarelhpeyykdc 237

QY 176 PEREVLWRFDLSRLAFHHVARELHPYFKNC 206

DB 207 pekeviewrfdsklathhvarelhpeyykdc 237

QY 176 PEREVLWRFDLSRLAFHHVARELHPYFKNC 206

DB 207 pekeviewrfdsklathhvarelhpeyykdc 237

QY 176 PEREVLWRFDLSRLAFHHVARELHPYFKNC 206

DB 207 pekeviewrfdsklathhvarelhpeyykdc 237

QY 176 PEREVLWRFDLSRLAFHHVARELHPYFKNC 206

DB 207 pekeviewrfdsklathhvarelhpeyykdc 237

QY 176 PEREVLWRFDLSRLAFHHVARELHPYFKNC 206

DB 207 pekeviewrfdsklathhvarelhpeyykdc 237

QY 176 PEREVLWRFDLSRLAFHHVARELHPYFKNC 206

DB 207 pekeviewrfdsklathhvarelhpeyykdc 237

QY 176 PEREVLWRFDLSRLAFHHVARELHPYFKNC 206

DB 207 pekeviewrfdsklathhvarelhpeyykdc 237

QY 176 PEREVLWRFDLSRLAFHHVARELHPYFKNC 206

DB 207 pekeviewrfdsklathhvarelhpeyykdc 237

QY 176 PEREVLWRFDLSRLAFHHVARELHPYFKNC 206

DB 207 pekeviewrfdsklathhvarelhpeyykdc 237

QY 176 PEREVLWRFDLSRLAFHHVARELHPYFKNC 206

DB 207 pekeviewrfdsklathhvarelhpeyykdc 237

QY 176 PEREVLWRFDLSRLAFHHVARELHPYFKNC 206

DB 207 pekeviewrfdsklathhvarelhpeyykdc 237

QY 176 PEREVLWRFDLSRLAFHHVARELHPYFKNC 206

DB 207 pekeviewrfdsklathhvarelhpeyykdc 237

QY 176 PEREVLWRFDLSRLAFHHVARELHPYFKNC 206

RESULT 29

AAB69363

ID AAB69363 standard; Protein; 206 AA.

XX AAB69363;

AC AAB69363;

XX AAB69363;

XX AAB69363;

XX AAB69363;

XX AAB69363;

XX AAB69363;

XX AAB69363;

XX AAB69363;

XX AAB69363;

XX AAB69363;

XX AAB69363;

XX AAB69363;

XX AAB69363;

XX AAB69363;

XX AAB69363;

XX AAB69363;

XX AAB69363;

XX AAB69363;

XX AAB69363;

XX AAB69363;

XX AAB69363;

DT 20-APR-2001 (first entry)

DE HIV-1 non-subtype B clone 962M651-8 nef protein.

XX HIV-1; human immunodeficiency virus; non-subtype B; gag; pol; env;

KW vpu; vif; vpr; tat; rev; nef; vaccine.

XX Human immunodeficiency virus type 1.

OS WO200026416-A1.

PN 11-MAY-2000.

XX 25-OCT-1999; 99WO-US24837.

XX 02-NOV-1998; 98US-0184418.

XX (UABR-) UAB RES FOUND.

XX Hahn BH, Shaw GM, Gao F;

XX

XX

XX

XX

XX

XX

XX

DR WPI: 2000-36561/31.  
XX  
XX Novel genomic nucleic acids of non-subtype B human immunodeficiency  
PT virus type 1 useful for detecting and treating AIDS comprises a  
PT specific nucleotide sequence  
XX  
XX Claim 41; Fig 22; 131pp; English.  
XX  
XX The present invention provides the protein and coding sequences for a  
CC number of human immunodeficiency virus (HIV) type 1 non-subtype B  
CC isolates. The sequences shown include the near full-length coding  
CC sequences from each isolate, and the env, pol, vif, vpr, gag, tat,  
CC rev and nef proteins. These can be used to detect the presence of HIV-1  
CC in a sample and to produce antibodies against non-subtype B HIV-1 virus.  
CC These antibodies can be used in vaccines to prevent and treat HIV  
CC infection.  
XX  
XX Sequence 206 AA;  
SQ  
  
Query Match 56.4%; Score 952; DB 21; Length 206;  
Best Local Similarity 81.1%; Pred. No. 2.1e-84;  
Matches 167; Conservative 24; Mismatches 15; Indels 0; Gaps 0;  
  
QY 1 MGCKWSSVVGWPTVRRMRRAEPAADGVGAASRDLEKHGALTSSNTAATNAACAWLEA 60  
Db 1 mgckwssvvgwptvrrmrrepaadgvgasqldkygaltssntstnaacawlea 60  
  
QY 61 QEEEEVGFVTPQVPLRPMYTKAAVDLSHFLKEKGLGGLHSQRQDILDLYIHTQGY 120  
Db 61 qeeeevgfvpqpvrpmtkykaavdlshflkekgleglihsqrqdiildlyihtqgy 120  
  
QY 121 FPDWQNTPGGVRYPLTFGWCYKLVPEPKVEEANKGENTSLHHPVSLHGMDDPREV 180  
Db 121 fpdwqntpggvrypltfgwcylvpdpkveeankgentslhlpvslhgmdddhrev 180  
  
QY 181 LEWRFSRLAFHHVARELHPEYFKNC 206  
Db 181 lkwrfsrlafhhvarelhpeyfknc 206  
  
RESULT 30  
AAB86169  
ID AAB86169 standard; Protein; 3025 AA.  
XX  
XX AAB86169;  
XX  
XX 21-AUG-2001 (first entry)  
DT  
DE HIV-1 subtype C protein fragment #1.  
DE  
XX  
XX Infection; diagnosis; human; humoral immune response; antiviral;  
KW cellular immune response; vaccine; treatment; gene therapy.  
KW  
XX  
XX Human immunodeficiency virus type 1.  
XX  
XX Key Location/Qualifiers  
FH Misc-difference 1..3025  
FT /label= Xaa  
FT /note= "Xaa represents a stop codon"  
XX  
XX DE10056747-A1.  
XX  
XX 31-MAY-2001.  
PD  
XX  
XX 16-NOV-2000; 2000DE-1056747.  
PF  
XX  
XX 16-NOV-1999; 99DE-1055089.  
PR  
XX  
XX (SHAO/) SHAO Y.  
PA (GENE-) GENEART GMBH GES ANGEWANDTE BIOTECHNOLOG.  
PA  
XX Wagner R, Wolf H, Shao Y, Graf M;  
PI

XX  
DR WPI: 2001-336417/36.  
DR N-PSDB; AAH20870.  
XX  
PT New nucleic acid sequences from a human immune deficiency virus  
PT intersubtype, useful for treatment, prevention and diagnosis of  
PT infection  
PT  
XX  
XX Disclosure; Fig 8A-O; 48pp; German.  
PS  
XX This invention describes a novel polynucleotide isolated from human  
CC immunodeficiency virus type 1 subtype C/B' which can be used for the  
CC induction of specific humoral and cellular immune responses. (I) and  
CC polypeptides (II) encoded by them, are useful in pharmaceuticals,  
CC vaccines and diagnostic agents, particularly for treatment or prevention  
CC of human immune deficiency virus-1 (HIV-1) infections, also for rational  
CC design of test or therapeutic reagents, or gene therapy vectors.  
CC Polypeptides, especially antibodies, specifically directed against (II)  
CC are similarly useful as pharmaceutical and diagnostic agents. (I) are  
CC specific for intersubtype C/B' of HIV-1 so are useful in regions  
CC (particularly China and South-East Asia) where this subtype is prevalent.  
CC The products of the invention have antiviral activity. This sequence  
CC represents a protein encoded by the HIV-1 subtype C genome described in  
CC the method of the invention.  
XX  
XX Sequence 3025 AA;  
SQ  
  
Query Match 56.2%; Score 948.5; DB 22; Length 3025;  
Best Local Similarity 70.6%; Pred. No. 1.9e-82;  
Matches 175; Conservative 22; Mismatches 42; Indels 9; Gaps 2;  
  
QY 1 MGCKWSSVVGWPTVRRMRRAEPAADGVGAASRDLEKHGALTSSNTAATNAACAWLEA 60  
Db 2723 mgckwssvvgwptvrrmrrepaadgvgasrdlekngaltssntaatnedcawlea 2782  
  
QY 61 QEEEEVGFVTPQVPLRPMYTKAAVDLSHFLKEKGLGGLHSQRQDILDLYIHTQGY 120  
Db 2783 qeeevgfvpqpvrpmtkykaavdlshflkekgleglihsqrqdiildlyihtqgy 2842  
  
QY 121 FPDWQNTPGGVRYPLTFGWCYKLVPEPKVEEANKGENTSLHHPVSLHGMDDPREV 180  
Db 2843 fpdwqntpggvrypltfgwcylvpdpkveeankgentslhlpvslhgmdddhrev 2902  
  
QY 181 LEWRFSRLAFHHVARELHPEYFKNCSTSEVPDPR-----LEPWKHPGSPKTTACTN 233  
Db 2903 lkwrfsrlafhhvarelhpeyfkncstsevpdpr-----lepwkhpqspkttactn 2960  
  
QY 234 CKKCCFHC 241  
Db 2961 ykqllfac 2968  
  
Search completed: August 26, 2002, 08:12:25  
Job time: 234 sec

GenCore version 4.5  
Copyright (c) 1993 - 2000 Compugen Ltd.

OM protein - protein search, using sw model

Run on: August 26, 2002, 08:08:31 ; Search time 38.46 Seconds  
(without alignments)  
191.797 Million cell updates/sec

Title: US-09-509-239-13  
Perfect score: 1688  
Sequence: 1 MGGKWSKSSVVGWPTVRM.....QSRGDPGPKTSGHHHHH 302

Scoring table: BLOSUM62  
Gapop 10.0 , Gapext 0.5

Searched: 231628 seqs, 24425594 residues

Total number of hits satisfying chosen parameters: 231628

Minimum DB seq length: 0  
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

Database : Issued\_Patents\_AA:\*  
1: /cgn2.6/ptodata/2/iaa/5A-COMB.pep.\*  
2: /cgn2.6/ptodata/2/iaa/5B-COMB.pep.\*  
3: /cgn2.6/ptodata/2/iaa/6A-COMB.pep.\*  
4: /cgn2.6/ptodata/2/iaa/6B-COMB.pep.\*  
5: /cgn2.6/ptodata/2/iaa/PTUS-COMB.pep.\*  
6: /cgn2.6/ptodata/2/iaa/backfiles1.pep.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	1138	67.4	239	4	US-08-679-493A-76
2	1115	66.1	206	3	US-08-388-353-644
3	1115	66.1	206	3	US-08-488-551B-644
4	971	57.5	206	4	US-08-679-493A-75
5	638	37.8	123	4	US-09-124-900-10
6	498.5	29.5	151	4	US-08-679-493A-74
7	486	28.8	86	1	US-08-450-257-1
8	486	28.8	86	1	US-08-450-246-1
9	486	28.8	86	1	US-08-450-098-1
10	486	28.8	86	1	US-08-451-233-1
11	486	28.8	86	1	US-08-450-236-1
12	486	28.8	86	3	US-07-808-452-1
13	486	28.8	86	3	US-09-030-613-19
14	486	28.8	86	4	US-09-124-900-6
15	486	28.8	86	4	US-09-451-905-19
16	486	28.8	86	4	US-08-235-403-1
17	486	28.8	86	5	PCR-US92-10770-1
18	486	28.8	86	5	PCR-US95-06077-2
19	478.5	28.3	266	3	US-08-815-809-5
20	478.5	28.3	280	2	US-08-816-155B-43
21	478.5	28.3	280	4	US-09-079-587-43
22	470	27.8	86	2	US-08-505-210-1
23	470	27.8	86	4	US-09-099-333-1
24	466	27.6	83	2	US-08-417-210A-96
25	455	27.0	82	1	US-08-053-079A-15
26	430.5	25.5	253	2	US-08-659-251-4
27	430.5	25.5	253	4	US-09-256-490-4

28	430.5	25.5	253	5	PCT-US96-11445-4	Sequence 4, Appl
29	409	24.2	72	3	US-09-030-613-17	Sequence 17, Appl
30	409	24.2	72	4	US-09-451-905-17	Sequence 17, Appl
31	385	22.8	72	2	US-08-893-853-1	Sequence 1, Appl
32	385	22.8	72	4	US-09-113-921-1	Sequence 1, Appl
33	372	22.0	122	6	5304466-3	Patent No. 5304466
34	365.5	21.7	312	1	US-08-094-128A-27	Sequence 27, Appl
35	365.5	21.7	312	1	US-08-455-674-27	Sequence 27, Appl
36	365.5	21.7	312	1	US-08-455-992-27	Sequence 27, Appl
37	365.5	21.7	312	1	US-08-455-972-27	Sequence 27, Appl
38	365.5	21.7	312	5	PCT-US92-00652-27	Sequence 27, Appl
39	362.5	21.5	3025	6	5223423-3	Patent No. 5223423
40	352	20.9	72	1	US-07-910-867B-2	Sequence 2, Appl
41	331	19.6	72	1	US-07-910-867B-5	Sequence 5, Appl
42	325	19.3	72	1	US-07-910-867B-3	Sequence 3, Appl
43	311	18.4	72	1	US-07-910-867B-1	Sequence 1, Appl
44	290	17.2	56	4	US-09-055-075C-48	Sequence 48, Appl
45	277	16.4	56	1	US-08-450-257-7	Sequence 7, Appl

ALIGNMENTS

RESULT 1  
US-08-679-493A-76  
; Sequence 76, Application US/08679493A  
; Patent No. 6303295  
; GENERAL INFORMATION:  
; APPLICANT: Taylor, Ethan W.  
; TITLE OF INVENTION: SELENOPROTEINS, CODING SEQUENCES AND METHODS  
; FILE REFERENCE: 55-95  
; CURRENT APPLICATION NUMBER: US/08/679,493A  
; CURRENT FILING DATE: 1996-07-12  
; PRIOR APPLICATION NUMBER: 60/001203  
; PRIOR FILING DATE: 1995-07-14  
; PRIOR APPLICATION NUMBER: 60/003,112  
; PRIOR FILING DATE: 1995-09-01  
; NUMBER OF SEQ ID NOS: 216  
; SOFTWARE: Patentin Ver. 2.0  
; SEQ ID NO 76  
; LENGTH: 239  
; TYPE: PRT  
; ORGANISM: Human immunodeficiency virus type 1  
; FEATURE:  
; NAME/KEY: VARIANT  
; LOCATION: (1)..(239)  
; OTHER INFORMATION: X os selenocysteine.  
US-08-679-493A-76

Query Match	67.4%	Score 1138;	DB 4;	Length 239;
Best Local Similarity	88.7%	Pred. No. 1.2e-108;		
Matches	211;	Conservative	2;	Mismatches 25; Indels 0; Gaps 0;
QY	1	MGGKWSKSSVVGWPTVRMRRAEPADGVGAASRDLKHKGAITSSNTAATNAACAWLEA	60	
Db	1	MGGKWSKSSVVGWPTVRMRRAEPADGVGAASRDLKHKGAITSSNTAATNAACAWLEA	60	
QY	61	QEEEEVGFPVTPQVPLRPMTYKAAVDLSHFLKKGGLIHSORQDIDLWYHTQGY	120	
Db	61	QEEEEVGFPVTPQVPLRPMTYKAAVDLSHFLKKGGLIHSORQDIDLWYHTQGY	120	
QY	121	FPDQNYTPGPGVRYPLTFGWCYKLVPEPDKVEEANKGENTSLHHPVSLHGMDPEREV	180	
Db	121	FPDQNYTPGPGVRYPLTFGWCYKLVPEPDKVEEANKGENTSLHHPVSLHGMDPEREV	180	
QY	181	LEWRFDSRLAFHVAHELHPEYFNKNTCTSEPDRLPELFWKHPGQPKTACTNCYCKKCC	238	
Db	181	LEWRFDSRLAFHVAHELHPEYFNKNTCTSEPDRLPELFWKHPGQPKTACTNCYCKKCC	238	

RESULT 2  
US-08-388-353-644

; Sequence 644, Application US/08388353  
; Patent No. 6010895  
; GENERAL INFORMATION:  
; APPLICANT: Deacon, Nicholas J.  
; APPLICANT: Learmont, Jennifer C.  
; APPLICANT: McPhee, Dale A.  
; APPLICANT: Crowe, Suzanne  
; APPLICANT: Cooper, David  
; TITLE OF INVENTION: NON-PATHOGENIC STRAINS OF HIV-1  
; NUMBER OF SEQUENCES: 800  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: Scully, Scott, Murphy & Presser  
; STREET: 400 Garden City Plaza  
; CITY: Garden City  
; STATE: New York  
; COUNTRY: United States  
; ZIP: 11530  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Floppy disk  
; COMPUTER: IBM PC compatible  
; OPERATING SYSTEM: PC-DOS/MS-DOS  
; SOFTWARE: Patent In Release #1.0, Version #1.25  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/08/388,353  
; FILING DATE: 14-FEB-1995  
; CLASSIFICATION: 424  
; ATTORNEY/AGENT INFORMATION:  
; NAME: Digiglio, Frank S.  
; REGISTRATION NUMBER: 31,346  
; REFERENCE/DOCKET NUMBER: 9606  
; TELECOMMUNICATION INFORMATION:  
; TELEPHONE: (516) 742-4343  
; TELEFAX: (516) 742-4366  
; TELEX: 230 901 SANS UR  
; INFORMATION FOR SEQ ID NO: 644:  
; SEQUENCE CHARACTERISTICS:  
; LENGTH: 206 amino acids  
; TYPE: amino acid  
; STRANDEDNESS: single  
; TOPOLOGY: linear  
; MOLECULE TYPE: protein  
; US-08-388-353-644

Query Match 66.1%; Score 1115; DB 3; Length 206;  
Best Local Similarity 98.1%; Pred. No. 2.le-106;  
Matches 202; Conservative 1; Mismatches 3; Indels 0; Gaps 0;  
  
QY 1 MGGKWSKSSVVGWPTVRRMRRAEPAADGVGAASRDLEKKGATSSNTAATNAACAWLEA 60  
Db 1 MGGKWSKSSVVGWPTVRRMRRAEPAADGVGAASRDLEKKGATSSNTAATNAACAWLEA 60  
  
QY 61 QEEEEVGFPVTPQVPLRPMTYKAAVDLSHFLKKEGGLGLHSORRODILDLYHTQGY 120  
Db 61 QEEEEVGFPVTPQVPLRPMTYKAAVDLSHFLKKEGGLGLHSORRODILDLYHTQGY 120  
  
QY 121 FPDQNTPTGPGVRYPLTFGWCYKLVPEPDKVEEANKGENTSLHVPVSLHGMDPPEV 180  
Db 121 FPDQNTPTGPGVRYPLTFGWCYKLVPEPDKVEEANKGENTSLHVPVSLHGMDPPEV 180  
  
QY 181 LEWRFDSRLAFHHVARELHPEYFKNC 206  
Db 181 LEWRFDSRLAFHHVARELHPEYFKNC 206

RESULT 3  
US-08-488-551B-644  
; Sequence 644, Application US/08488551B  
; Patent No. 6015661  
; GENERAL INFORMATION:  
; APPLICANT: Nicholas J. Deacon  
; APPLICANT: Dale A. McPhee  
; APPLICANT: David Cooper

; TITLE OF INVENTION: NON-PATHOGENIC STRAINS OF HIV-1  
; NUMBER OF SEQUENCES: 841  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: Scully, Scott, Murphy & Presser  
; STREET: 400 GARDEN CITY PLAZA  
; CITY: GARDEN CITY  
; STATE: NEW YORK  
; COUNTRY: U.S.A.  
; ZIP: 11530-0299  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Floppy disk  
; COMPUTER: IBM PC compatible  
; OPERATING SYSTEM: PC-DOS/MS-DOS  
; SOFTWARE: Patent In Release #1.0, Version #1.25  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/08/488,551B  
; FILING DATE: 07-JUN-1995  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: PM3864 (AU)  
; FILING DATE: 14-FEB-1994  
; APPLICATION NUMBER: PM4002 (AU)  
; FILING DATE: 21-FEB-1994  
; APPLICATION NUMBER: PM0284 (AU)  
; FILING DATE: 23-DEC-1994  
; APPLICATION NUMBER: US 08/388,353  
; FILING DATE: 14-FEB-1995  
; APPLICATION NUMBER: PM3021/95  
; FILING DATE: 17-MAY-1995  
; ATTORNEY/AGENT INFORMATION:  
; NAME: FRANK S. DIGIGLIO  
; REFERENCE/DOCKET NUMBER: 9606Z  
; TELECOMMUNICATION INFORMATION:  
; TELEPHONE: (516) 742-4343  
; TELEFAX: (516) 742-4366  
; INFORMATION FOR SEQ ID NO: 644:  
; SEQUENCE CHARACTERISTICS:  
; LENGTH: 206 amino acids  
; TYPE: amino acid  
; STRANDEDNESS: single  
; TOPOLOGY: linear  
; MOLECULE TYPE: protein  
; US-08-488-551B-644

Query Match 66.1%; Score 1115; DB 3; Length 206;  
Best Local Similarity 98.1%; Pred. No. 2.le-106;  
Matches 202; Conservative 1; Mismatches 3; Indels 0; Gaps 0;  
  
QY 1 MGGKWSKSSVVGWPTVRRMRRAEPAADGVGAASRDLEKKGATSSNTAATNAACAWLEA 60  
Db 1 MGGKWSKSSVVGWPTVRRMRRAEPAADGVGAASRDLEKKGATSSNTAATNAACAWLEA 60  
  
QY 61 QEEEEVGFPVTPQVPLRPMTYKAAVDLSHFLKKEGGLGLHSORRODILDLYHTQGY 120  
Db 61 QEEEEVGFPVTPQVPLRPMTYKAAVDLSHFLKKEGGLGLHSORRODILDLYHTQGY 120  
  
QY 121 FPDQNTPTGPGVRYPLTFGWCYKLVPEPDKVEEANKGENTSLHVPVSLHGMDPPEV 180  
Db 121 FPDQNTPTGPGVRYPLTFGWCYKLVPEPDKVEEANKGENTSLHVPVSLHGMDPPEV 180  
  
QY 181 LEWRFDSRLAFHHVARELHPEYFKNC 206  
Db 181 LEWRFDSRLAFHHVARELHPEYFKNC 206

RESULT 4  
US-08-679-493A-75  
; Sequence 75, Application US/08679493A  
; Patent No. 6303295  
; GENERAL INFORMATION:  
; APPLICANT: Taylor, Ethan W.  
; TITLE OF INVENTION: SELENOPROTEINS, CODING SEQUENCES AND METHODS  
; FILE REFERENCE: 55-95

;; CURRENT APPLICATION NUMBER: US/08/679,493A  
;; CURRENT FILING DATE: 1996-07-12  
;; PRIOR APPLICATION NUMBER: 60/001203  
;; PRIOR FILING DATE: 1995-07-14  
;; PRIOR APPLICATION NUMBER: 60/003,112  
;; PRIOR FILING DATE: 1995-09-01  
;; NUMBER OF SEQ ID NOS: 216  
;; SOFTWARE: PatentIn Ver. 2.0  
;; SEQ ID NO 75  
;; LENGTH: 206  
;; TYPE: PRT  
;; ORGANISM: Human immunodeficiency virus type 1  
;; NAME/KEY: VARIANT  
;; LOCATION: (1)..(206)  
;; OTHER INFORMATION: X is selenocysteine.  
US-08-679-493A-75

Query Match 57.5%; Score 971; DB 4; Length 206;  
Best Local Similarity 89.0%; Pred. No. 1.2e-91;  
Matches 186; Conservative 2; Mismatches 15; Indels 6; Gaps 3;  
QY 1 MGCKWSKSSVVGWPTVRMRRAEPAADGVGAASRDLEKHGAISSNTAATNAACAWLEA 60  
Db 1 MGCKWSKSSVVGWPTVRMRRAEPAADGVGAASRDLEKHGAISSNTAATNAACAWLEA 60  
QY 61 QEEVEVGFPTPOVPLRPMTYKAADVLSHFLKEKGG---LEGLIHSORRODILDLWIYHT 117  
Db 61 QEEVEVGFPTPOVPLRPMTYKAADVLSHFLKEKGG---LEGLIHSORRODILDLWIYHT 117  
QY 118 QGVFPDQNTPGGVRYPLTFCGCKVLPVPEPKVEEANKGENTSLHHPVSLHGMDDPE 177  
Db 119 R-LLPQNTPGGVRYPLTFCGCKVLPVPEPKVEEANKGENTSLHHPVSLHGMDDPE 177  
QY 178 REVLEWRFDSRLAFHHVARELHPEYFKNC 206  
Db 178 REVLEWRFDSRLAFHHVARELHPEYFKNC 206

RESULT 5  
US-09-124-900-10  
;; Sequence 10, Application US/09124900  
;; Patent No. 6268484  
;; GENERAL INFORMATION:  
;; APPLICANT: KATZINGER, Hermann  
;; APPLICANT: BUCHACHER, Andrea  
;; APPLICANT: ERNST, Wolfgang  
;; APPLICANT: BALLAUN, Claudia  
;; APPLICANT: PURTSCHER, Martin  
;; APPLICANT: TRKOLA, Alexandra  
;; APPLICANT: PREDL, Renate  
;; APPLICANT: SCHMATZ, Christine  
;; APPLICANT: KLIMA, Annelies  
;; APPLICANT: STEINDL, Franz  
;; APPLICANT: MUSTER, Thomas  
;; TITLE OF INVENTION: HIV-Vaccines  
;; FILE REFERENCE: 1939-112P  
;; CURRENT APPLICATION NUMBER: US/09/124,900  
;; CURRENT FILING DATE: 1998-07-30  
;; PRIOR APPLICATION NUMBER: PCT/EP95/01481  
;; PRIOR FILING DATE: 1995-04-19  
;; NUMBER OF SEQ ID NOS: 11  
;; SOFTWARE: PatentIn version 3.0  
;; SEQ ID NO 10  
;; LENGTH: 123  
;; TYPE: PRT  
;; ORGANISM: Human immunodeficiency virus type 1  
US-09-124-900-10

Query Match 37.8%; Score 638; DB 4; Length 123;  
Best Local Similarity 97.6%; Pred. No. 7.3e-58;

Matches 120; Conservative 0; Mismatches 3; Indels 0; Gaps 0;  
QY 1 MGCKWSKSSVVGWPTVRMRRAEPAADGVGAASRDLEKHGAISSNTAATNAACAWLEA 60  
Db 1 MGCKWSKSSVVGWPTVRMRRAEPAADGVGAASRDLEKHGAISSNTAATNAACAWLEA 60  
QY 61 QEEVEVGFPTPOVPLRPMTYKAADVLSHFLKEKGGLEGLIHSORRODILDLWIYHTGY 120  
Db 61 QEEVEVGFPTPOVPLRPMTYKAADVLSHFLKEKGGLEGLIHSORRODILDLWIYHTGY 120  
QY 121 FPD 123  
Db 121 FPD 123  
RESULT 6  
US-08-679-493A-74  
;; Sequence 74, Application US/08679493A  
;; Patent No. 6303295  
;; GENERAL INFORMATION:  
;; APPLICANT: Taylor, Ethan W.  
;; TITLE OF INVENTION: SELENOPROTEINS, CODING SEQUENCES AND METHODS  
;; FILE REFERENCE: 55-95  
;; CURRENT APPLICATION NUMBER: US/08/679,493A  
;; CURRENT FILING DATE: 1996-07-12  
;; PRIOR APPLICATION NUMBER: 60/001203  
;; PRIOR FILING DATE: 1995-07-14  
;; PRIOR APPLICATION NUMBER: 60/003,112  
;; PRIOR FILING DATE: 1995-09-01  
;; NUMBER OF SEQ ID NOS: 216  
;; SOFTWARE: PatentIn Ver. 2.0  
;; SEQ ID NO 74  
;; LENGTH: 151  
;; TYPE: PRT  
;; ORGANISM: Human immunodeficiency virus type 1  
;; FEATURE:  
;; NAME/KEY: VARIANT  
;; LOCATION: (1)..(151)  
;; OTHER INFORMATION: X is selenocysteine.  
US-08-679-493A-74

Query Match 29.5%; Score 498.5; DB 4; Length 151;  
Best Local Similarity 73.0%; Pred. No. 1.8e-43;  
Matches 103; Conservative 4; Mismatches 7; Indels 27; Gaps 3;  
QY 1 MGCKWSKSSVVGWPTVRMRRAEPAADGVGAASRDLEKHGAISSNTAATNAACAWLEA 60  
Db 1 MGCKWSKSSVVGWPTVRMRRAEPAADGVGAASRDLEKHGAISSNTAATNAACAWLEA 60  
QY 61 QEEVEVGFPTPOVPLRPMTYKAADVLSHFLKEKGG----- 96  
Db 61 QEEVEVGFPTPOVPLRPMTYKAADVLSHFLKEKGG----- 96  
QY 97 LEGL--IHSORRODI-LDLWI 114  
Db 121 LPXLAEHLTRARGISTDLWM 141

RESULT 7  
US-08-450-257-1  
;; Sequence 1, Application US/08450257  
;; Patent No. 5652122  
;; GENERAL INFORMATION:  
;; APPLICANT: FRANKEL, Alan  
;; APPLICANT: PABO, Carl  
;; APPLICANT: BARSOUM, James G.  
;; APPLICANT: FAWELL, Stephen E.  
;; APPLICANT: PEPINSKY, R. B.  
;; TITLE OF INVENTION: TAT-DERIVED TRANSPORT POLYPEPTIDES  
;; NUMBER OF SEQUENCES: 69  
;; CORRESPONDENCE ADDRESS:  
;; ADDRESSEE: FISH & NEAVE

STREET: 1251 Avenue of the Americas  
CITY: New York  
STATE: New York  
COUNTRY: USA  
ZIP: 10020  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: Patent in Release #1.0, Version #1.25  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/450.257  
FILING DATE: 25-MAY-1995  
CLASSIFICATION: 514  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: US 08/235.403  
FILING DATE: 28-APR-1994  
APPLICATION NUMBER: US 07/934.375  
FILING DATE: 21-AUG-1992  
APPLICATION NUMBER: US 07/098.766  
FILING DATE: 28-JUL-1993  
APPLICATION NUMBER: PCT/US93/07833  
FILING DATE: 19-AUG-1993  
APPLICATION NUMBER: US 07/454.450  
FILING DATE: 21-DEC-1989  
APPLICATION NUMBER: US 07/636.662  
FILING DATE: 02-JAN-1991  
APPLICATION NUMBER: US 08/158.015  
FILING DATE: 24-NOV-1993  
ATTORNEY/AGENT INFORMATION:  
NAME: Haley Jr., James F.  
REGISTRATION NUMBER: 27,794  
REFERENCE/DOCKET NUMBER: B170 CIP 2  
TELEPHONE: (212) 596-9000  
TELEFAX: (212) 596-9090  
TELEX: 14-8367  
INFORMATION FOR SEQ ID NO: 1:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 86 amino acids  
TYPE: amino acid  
TOPOLOGY: linear  
MOLECULE TYPE: protein  
ORIGINAL SOURCE:  
ORGANISM: human immunodeficiency virus  
STRAIN: type 1  
US-08-450-257-1

Query Match 28.8%; Score 486; DB 1; Length 86;  
Best Local Similarity 100.0%; Pred. No. 1.6e-42;  
Matches 85; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
QY 209 EPVDPRLPEWKHPGSPKPTACTNCYCKKCCFHCQVCFITKALGISYGRKKRRRRPPQG 268  
DB 2 EPVDPRLPEWKHPGSPKPTACTNCYCKKCCFHCQVCFITKALGISYGRKKRRRRPPQG 61  
QY 269 SOTHQVSLSKQPTSQSRGDPGPK 293  
DB 62 SOTHQVSLSKQPTSQSRGDPGPK 86

RESULT 8  
US-08-450-246-1  
Sequence 1, Application US/08450246  
Patent No. 5670617  
GENERAL INFORMATION:  
APPLICANT: FRANKEL, Alan  
APPLICANT: BARSOUM, James G.  
APPLICANT: FAMELL, Stephen E.  
APPLICANT: PEPINSKY, R. B.  
TITLE OF INVENTION: TAT-DERIVED TRANSPORT POLYPEPTIDES

NUMBER OF SEQUENCES: 69  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: FISH & NEAVE  
STREET: 1251 Avenue of the Americas  
CITY: New York  
STATE: New York  
COUNTRY: USA  
ZIP: 10020  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: Patent in Release #1.0, Version #1.25  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/450.246  
FILING DATE: 25-MAY-1995  
CLASSIFICATION: 514  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: US 08/235.403  
FILING DATE: 28-APR-1994  
APPLICATION NUMBER: US 07/934.375  
FILING DATE: 21-AUG-1992  
APPLICATION NUMBER: US 07/098.766  
FILING DATE: 28-JUL-1993  
APPLICATION NUMBER: PCT/US93/07833  
FILING DATE: 19-AUG-1993  
APPLICATION NUMBER: US 07/454.450  
FILING DATE: 21-DEC-1989  
APPLICATION NUMBER: US 07/636.662  
FILING DATE: 02-JAN-1991  
APPLICATION NUMBER: US 08/158.015  
FILING DATE: 24-NOV-1993  
ATTORNEY/AGENT INFORMATION:  
NAME: Haley Jr., James F.  
REGISTRATION NUMBER: 27,794  
REFERENCE/DOCKET NUMBER: B170 CIP 2  
TELEPHONE: (212) 596-9000  
TELEFAX: (212) 596-9090  
TELEX: 14-8367  
INFORMATION FOR SEQ ID NO: 1:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 86 amino acids  
TYPE: amino acid  
TOPOLOGY: linear  
MOLECULE TYPE: protein  
ORIGINAL SOURCE:  
ORGANISM: human immunodeficiency virus  
STRAIN: type 1  
US-08-450-246-1  
Query Match 28.8%; Score 486; DB 1; Length 86;  
Best Local Similarity 100.0%; Pred. No. 1.6e-42;  
Matches 85; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
QY 209 EPVDPRLPEWKHPGSPKPTACTNCYCKKCCFHCQVCFITKALGISYGRKKRRRRPPQG 268  
DB 2 EPVDPRLPEWKHPGSPKPTACTNCYCKKCCFHCQVCFITKALGISYGRKKRRRRPPQG 61  
QY 269 SOTHQVSLSKQPTSQSRGDPGPK 293  
DB 62 SOTHQVSLSKQPTSQSRGDPGPK 86  
RESULT 9  
US-08-450-098-1  
Sequence 1, Application US/08450098  
Patent No. 5674980  
GENERAL INFORMATION:  
APPLICANT: FRANKEL, Alan  
APPLICANT: BARSOUM, James G.  
TITLE OF INVENTION: TAT-DERIVED TRANSPORT POLYPEPTIDES

APPLICANT: FAWELL, Stephen E.  
APPLICANT: PEPINSKY, R. B.  
TITLE OF INVENTION: TAT-DERIVED TRANSPORT POLYPEPTIDES  
NUMBER OF SEQUENCES: 69  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: FISH & NEAVE  
STREET: 1251 Avenue of the Americas  
CITY: New York  
STATE: New York  
COUNTRY: USA  
ZIP: 10020  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: Patent in Release #1.0, Version #1.25  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/450,098  
FILING DATE: 25-MAY-1995  
CLASSIFICATION: 514  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: US 08/235,403  
FILING DATE: 28-APR-1994  
APPLICATION NUMBER: US 07/934,375  
FILING DATE: 21-AUG-1992  
APPLICATION NUMBER: US 07/098,766  
FILING DATE: 28-JUL-1993  
APPLICATION NUMBER: PCT/US93/07833  
FILING DATE: 19-AUG-1993  
APPLICATION NUMBER: US 07/454,450  
FILING DATE: 21-DEC-1989  
APPLICATION NUMBER: US 07/636,662  
FILING DATE: 02-JAN-1991  
APPLICATION NUMBER: US 08/158,015  
FILING DATE: 24-NOV-1993  
ATTORNEY/AGENT INFORMATION:  
NAME: Haley Jr., James F.  
REGISTRATION NUMBER: 27,794  
REFERENCE/DOCKET NUMBER: B170 CIP 2  
TELEPHONE: (212) 596-9000  
TELEFAX: (212) 596-9090  
TELEX: 14-8367  
INFORMATION FOR SEQ ID NO: 1:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 86 amino acids  
TYPE: amino acid  
TOPOLOGY: linear  
MOLECULE TYPE: protein  
ORIGINAL SOURCE:  
ORGANISM: human immunodeficiency virus  
STRAIN: type 1  
US-08-450-098-1

Query Match 28.8%; Score 486; DB 1; Length 86;  
Best Local Similarity 100.0%; Pred. No. 1.6e-42;  
Matches 85; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 209 EPVDPRLPEPKHFGSQPKTACTNCYCKKCCFHQVCFITKALGISYGRKKRRRRPPQG 268  
|||||  
Db 2 EPVDPRLPEPKHFGSQPKTACTNCYCKKCCFHQVCFITKALGISYGRKKRRRRPPQG 61

QY 269 SQTHQVSLSKQPTSQSRGDPGPK 293  
|||||  
Db 62 SQTHQVSLSKQPTSQSRGDPGPK 86

RESULT 10  
US-08-451-233-1  
; Sequence 1, Application US/08451233  
; Patent No. 5747641  
; GENERAL INFORMATION:

APPLICANT: FRANKEL, Alan  
APPLICANT: PABO, Carl  
APPLICANT: BARSOUM, James G.  
APPLICANT: FAWELL, Stephen E.  
APPLICANT: PEPINSKY, R. B.  
TITLE OF INVENTION: TAT-DERIVED TRANSPORT POLYPEPTIDES  
NUMBER OF SEQUENCES: 69  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: FISH & NEAVE  
STREET: 1251 Avenue of the Americas  
CITY: New York  
STATE: New York  
COUNTRY: USA  
ZIP: 10020  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: Patent in Release #1.0, Version #1.25  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/451,233  
FILING DATE: 25-MAY-1995  
CLASSIFICATION: 514  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: US 08/235,403  
FILING DATE: 28-APR-1994  
APPLICATION NUMBER: US 07/934,375  
FILING DATE: 21-AUG-1992  
APPLICATION NUMBER: US 07/098,766  
FILING DATE: 28-JUL-1993  
APPLICATION NUMBER: PCT/US93/07833  
FILING DATE: 19-AUG-1993  
APPLICATION NUMBER: US 07/454,450  
FILING DATE: 21-DEC-1989  
APPLICATION NUMBER: US 07/636,662  
FILING DATE: 02-JAN-1991  
APPLICATION NUMBER: US 08/158,015  
FILING DATE: 24-NOV-1993  
ATTORNEY/AGENT INFORMATION:  
NAME: Haley Jr., James F.  
REGISTRATION NUMBER: 27,794  
REFERENCE/DOCKET NUMBER: B170 CIP 2  
TELEPHONE: (212) 596-9000  
TELEFAX: (212) 596-9090  
TELEX: 14-8367  
INFORMATION FOR SEQ ID NO: 1:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 86 amino acids  
TYPE: amino acid  
TOPOLOGY: linear  
MOLECULE TYPE: protein  
ORIGINAL SOURCE:  
ORGANISM: human immunodeficiency virus  
STRAIN: type 1  
US-08-451-233-1

Query Match 28.8%; Score 486; DB 1; Length 86;  
Best Local Similarity 100.0%; Pred. No. 1.6e-42;  
Matches 85; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 209 EPVDPRLPEPKHFGSQPKTACTNCYCKKCCFHQVCFITKALGISYGRKKRRRRPPQG 268  
|||||  
Db 2 EPVDPRLPEPKHFGSQPKTACTNCYCKKCCFHQVCFITKALGISYGRKKRRRRPPQG 61

QY 269 SQTHQVSLSKQPTSQSRGDPGPK 293  
|||||  
Db 62 SQTHQVSLSKQPTSQSRGDPGPK 86

RESULT 11  
US-08-450-236-1



Sequence 1, Application US/08450236  
Patent No. 5804604  
GENERAL INFORMATION:  
APPLICANT: FRANKEL, Alan  
APPLICANT: PABO, Carl  
APPLICANT: BARSOUM, James G.  
APPLICANT: FAWELL, Stephen E.  
APPLICANT: PEPINSKY, R. B.  
TITLE OF INVENTION: TAT-DERIVED TRANSPORT POLYPEPTIDES  
NUMBER OF SEQUENCES: 69  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: FISH & NEAVE  
STREET: 1251 Avenue of the Americas  
CITY: New York  
STATE: New York  
COUNTRY: USA  
ZIP: 10020  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: Patent In Release #1.0, Version #1.25  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/450,236  
FILING DATE: 25-MAY-1995  
CLASSIFICATION: 514  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: US 08/235,403  
FILING DATE: 28-APR-1994  
APPLICATION NUMBER: US 07/934,375  
FILING DATE: 21-AUG-1992  
APPLICATION NUMBER: US 07/098,766  
FILING DATE: 28-JUL-1993  
APPLICATION NUMBER: PCT/US93/07833  
FILING DATE: 19-AUG-1993  
APPLICATION NUMBER: US 07/454,450  
FILING DATE: 21-DEC-1989  
APPLICATION NUMBER: US 07/636,662  
FILING DATE: 02-JAN-1991  
APPLICATION NUMBER: US 08/158,015  
FILING DATE: 24-NOV-1993  
ATTORNEY/AGENT INFORMATION:  
NAME: Haley Jr., James F.  
REGISTRATION NUMBER: 27,794  
REFERENCE/DOCKET NUMBER: B170 CIP 2  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: (212) 596-9000  
TELEFAX: (212) 596-9090  
TELEX: 14-8367  
INFORMATION FOR SEQ ID NO: 1:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 86 amino acids  
TYPE: amino acid  
TOPOLOGY: linear  
MOLECULE TYPE: protein  
ORIGINAL SOURCE:  
ORGANISM: human immunodeficiency virus  
STRAIN: type 1  
US-08-450-236-1

Query Match 28.8%; Score 486; DB 1; Length 86;  
Best Local Similarity 100.0%; Pred. No. 1.6e-42; Mismatches 0; Indels 0; Gaps 0;  
Matches 85; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
QY 209 EPVDPRLPEWKHPGSPKPTACTNCYCKKCFHCQVCFITKALGISYGRKKRRRRPPQG 268  
Db 2 EPVDPRLPEWKHPGSPKPTACTNCYCKKCFHCQVCFITKALGISYGRKKRRRRPPQG 61  
QY 269 SQTHQVLSKQPTSQSRGDPGPK 293  
Db 62 SQTHQVLSKQPTSQSRGDPGPK 86

RESULT 12  
US-07-808-452-1  
Sequence 1, Application US/07808452  
Patent No. 6063612  
GENERAL INFORMATION:  
APPLICANT: Jayasena, Sumedha D.  
APPLICANT: Johnston, Brian H.  
TITLE OF INVENTION: Antiviral Reagents Based on  
TITLE OF INVENTION: RNA-Binding Proteins  
NUMBER OF SEQUENCES: 24  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: Law Offices of Peter Dehlinger  
STREET: 350 Cambridge Avenue, Suite 300  
CITY: Palo Alto  
STATE: CA  
COUNTRY: USA  
ZIP: 94306  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: Patent In Release #1.0, Version #1.25  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/07/808,452  
FILING DATE: 19911213  
CLASSIFICATION: 514  
ATTORNEY/AGENT INFORMATION:  
NAME: Fabian, Gary R.  
REGISTRATION NUMBER: 33,875  
REFERENCE/DOCKET NUMBER: 8255-008  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: (415) 324-0880  
TELEFAX: (415) 324-0960  
INFORMATION FOR SEQ ID NO: 1:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 86 amino acids  
TYPE: AMINO ACID  
TOPOLOGY: linear  
MOLECULE TYPE: protein  
HYPOTHETICAL: NO  
ANTI-SENSE: NO  
ORIGINAL SOURCE:  
INDIVIDUAL ISOLATE: the sequence of the TAT protein of  
INDIVIDUAL ISOLATE: HIV-1  
US-07-808-452-1  
Query Match 28.8%; Score 486; DB 3; Length 86;  
Best Local Similarity 100.0%; Pred. No. 1.6e-42; Mismatches 0; Indels 0; Gaps 0;  
Matches 85; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
QY 209 EPVDPRLPEWKHPGSPKPTACTNCYCKKCFHCQVCFITKALGISYGRKKRRRRPPQG 268  
Db 2 EPVDPRLPEWKHPGSPKPTACTNCYCKKCFHCQVCFITKALGISYGRKKRRRRPPQG 61  
QY 269 SQTHQVLSKQPTSQSRGDPGPK 293  
Db 62 SQTHQVLSKQPTSQSRGDPGPK 86  
RESULT 13  
US-09-030-613-19  
Sequence 19, Application US/09030613  
Patent No. 6083706  
GENERAL INFORMATION:  
APPLICANT: Florikiewicz, Robert Z.  
APPLICANT: Baird, J. Andrew  
TITLE OF INVENTION: INHIBITORS OF LEADERLESS PROTEIN EXPORT  
NUMBER OF SEQUENCES: 36  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: SEED and BERRY LLP  
STREET: 6300 Columbia Center, 701 Fifth Avenue

```
; CITY: Seattle
; STATE: Washington
; COUNTRY: USA
; ZIP: 98104
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.30
; CURRENT APPLICATION NUMBER: US/09/030.613
; FILING DATE: 25-FEB-1998
; CLASSIFICATION:
; ATTORNEY/AGENT INFORMATION:
; NAME: No. 6083706tenburg Ph.D., Carol
; REGISTRATION NUMBER: 39,317
; REFERENCE/DOCKET NUMBER: 760100.418C1
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (206) 622-4900
; TELEFAX: (206) 682-6031
; INFORMATION FOR SEQ ID NO: 19:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 86 amino acids
; TYPE: amino acid
; STRANDEDNESS:
; TOPOLOGY: linear
US-09-030-613-19

Query Match 28.8%; Score 486; DB 3; Length 86;
Best Local Similarity 100.0%; Pred. No. 1.6e-42;
Matches 85; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 209 EPVDPRLPEPKHGPSQPKTACTNCYCKKCCFHCQVCFTKALGISYGRKKRRRPPQG 268
Db 2 EPVDPRLPEPKHGPSQPKTACTNCYCKKCCFHCQVCFTKALGISYGRKKRRRPPQG 61

QY 269 SOTHQVSLSKOPTSQRGDPGPK 293
Db 62 SOTHQVSLSKOPTSQRGDPGPK 86

RESULT 14
US-09-124-900-6
; Sequence 6, Application US/09124900
; Patent No. 6268484
; GENERAL INFORMATION:
; APPLICANT: KATINGER, Hermann
; APPLICANT: BUCHACHER, Andrea
; APPLICANT: ERNST, Wolfgang
; APPLICANT: BALLAUN, Claudia
; APPLICANT: PURTSCHER, Martin
; APPLICANT: TRKOLA, Alexandra
; APPLICANT: PREDL, Renate
; APPLICANT: SCHWATZ, Christine
; APPLICANT: KLIMA, Annelies
; APPLICANT: STEINDL, Franz
; APPLICANT: MUSTER, Thomas
; TITLE OF INVENTION: HIV-vaccines
; FILE REFERENCE: 1939-112P
; CURRENT APPLICATION NUMBER: US/09/124,900
; PRIOR FILING DATE: 1998-07-30
; PRIOR APPLICATION NUMBER: PCT/EP95/01481
; PRIOR FILING DATE: 1995-04-19
; NUMBER OF SEQ ID NOS: 11
; SOFTWARE: PatentIn version 3.0
; SEQ ID NO 6
; LENGTH: 86
; TYPE: PRT
; ORGANISM: Human immunodeficiency virus type 1
US-09-124-900-6
```

```
Query Match 28.8%; Score 486; DB 4; Length 86;
Best Local Similarity 100.0%; Pred. No. 1.6e-42;
Matches 85; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 209 EPVDPRLPEPKHGPSQPKTACTNCYCKKCCFHCQVCFTKALGISYGRKKRRRPPQG 268
Db 2 EPVDPRLPEPKHGPSQPKTACTNCYCKKCCFHCQVCFTKALGISYGRKKRRRPPQG 61

QY 269 SOTHQVSLSKOPTSQRGDPGPK 293
Db 62 SOTHQVSLSKOPTSQRGDPGPK 86

RESULT 15
US-09-451-905-19
; Sequence 19, Application US/09451905
; Patent No. 6306613
; GENERAL INFORMATION:
; APPLICANT: Robert Z. Florkiewicz
; APPLICANT: Andrew Baird
; APPLICANT: Dale E. Warnock
; TITLE OF INVENTION: MODULATORS OF LEADERLESS PROTEIN EXPORT
; FILE REFERENCE: 200124.402C4
; CURRENT APPLICATION NUMBER: US/09/451,905
; CURRENT FILING DATE: 1999-12-01
; NUMBER OF SEQ ID NOS: 48
; SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 19
; LENGTH: 86
; TYPE: PRT
; ORGANISM: Homo sapien
US-09-451-905-19

Query Match 28.8%; Score 486; DB 4; Length 86;
Best Local Similarity 100.0%; Pred. No. 1.6e-42;
Matches 85; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 209 EPVDPRLPEPKHGPSQPKTACTNCYCKKCCFHCQVCFTKALGISYGRKKRRRPPQG 268
Db 2 EPVDPRLPEPKHGPSQPKTACTNCYCKKCCFHCQVCFTKALGISYGRKKRRRPPQG 61

QY 269 SOTHQVSLSKOPTSQRGDPGPK 293
Db 62 SOTHQVSLSKOPTSQRGDPGPK 86

RESULT 16
US-08-235-403-1
; Sequence 1, Application US/08235403
; Patent No. 6316003
; GENERAL INFORMATION:
; APPLICANT: FRANKEL, Alan
; APPLICANT: PABO, Carl
; APPLICANT: BARSOU, James G.
; APPLICANT: FAWELL, Stephen E.
; APPLICANT: PEPINSKY, R. B.
; TITLE OF INVENTION: TAT-DERIVED TRANSPORT POLYPEPTIDES
; NUMBER OF SEQUENCES: 69
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: FISH & NEAVE
; STREET: 1251 Avenue of the Americas
; CITY: New York
; STATE: New York
; COUNTRY: USA
; ZIP: 10020
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
```

APPLICATION NUMBER: US/08/235.403  
FILING DATE: 13-DEC-1991  
CLASSIFICATION: 514  
PRIOR APPLICATION NUMBER: US 07/934.375  
FILING DATE: 21-AUG-1992  
APPLICATION NUMBER: US 07/098.766  
FILING DATE: 28-JUL-1993  
APPLICATION NUMBER: PCT/US93/07833  
FILING DATE: 19-AUG-1993  
APPLICATION NUMBER: US 07/454.450  
FILING DATE: 21-DEC-1989  
APPLICATION NUMBER: US 07/636.662  
FILING DATE: 02-JAN-1991  
APPLICATION NUMBER: US 08/158.015  
FILING DATE: 24-NOV-1993  
ATTORNEY/AGENT INFORMATION:  
NAME: Haley Jr., James F.  
REGISTRATION NUMBER: 27,794  
REFERENCE/DOCKET NUMBER: B170 CIP 2  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: (212) 596-9090  
TELEX: 14-8367  
INFORMATION FOR SEQ ID NO: 1:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 86 amino acids  
TYPE: amino acid  
TOPOLOGY: linear  
MOLECULE TYPE: protein  
ORIGINAL SOURCE:  
ORGANISM: human immunodeficiency virus  
STRAIN: type 1  
US-08-235-403-1

Query Match 28.8%; Score 486; DB 4; Length 86;  
Best Local Similarity 100.0%; Pred. No. 1.6e-42;  
Matches 85; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 209 EPVDPRLPWPVKHPSQPKTACTNCYKCCFHCQVCFITKALGISYGRKKRRRPPQG 268  
|||||  
Db 2 EPVDPRLPWPVKHPSQPKTACTNCYKCCFHCQVCFITKALGISYGRKKRRRPPQG 61

QY 269 SOTHQVSLSKOPTSOSRGDPTGPK 293  
|||||  
Db 62 SOTHQVSLSKOPTSOSRGDPTGPK 86

RESULT 17  
PCT-US92-10770-1  
Sequence 1, Application PC/TUS9210770  
GENERAL INFORMATION:  
APPLICANT: Jayasena, Suredha D.  
APPLICANT: Johnston, Brian H.  
TITLE OF INVENTION: Antiviral Reagents Based on  
TITLE OF INVENTION: RNA-Binding Proteins  
NUMBER OF SEQUENCES: 24  
CORRESPONDENCE ADDRESS:  
ADDRESS: 333 Ravenswood Avenue  
CITY: Menlo Park  
STATE: CA  
COUNTRY: USA  
ZIP: 94025  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: PatentIn Release #1.0, Version #1.25  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: PCT/US92/10770  
FILING DATE: 19921211

CLASSIFICATION:  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: US 07/808.452  
FILING DATE: 13-DEC-1991  
ATTORNEY/AGENT INFORMATION:  
NAME: Fabian, Gary R.  
REGISTRATION NUMBER: 33,875  
REFERENCE/DOCKET NUMBER: P-2962  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: (415) 859-4550  
TELEFAX: (415) 859-3880  
INFORMATION FOR SEQ ID NO: 1:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 86 amino acids  
TYPE: AMINO ACID  
TOPOLOGY: linear  
MOLECULE TYPE: protein  
HYPOTHETICAL: NO  
ANTI-SENSE: NO  
ORIGINAL SOURCE:  
INDIVIDUAL ISOLATE: the sequence of the TAT protein of  
INDIVIDUAL ISOLATE: HIV-1  
PCT-US92-10770-1

Query Match 28.8%; Score 486; DB 5; Length 86;  
Best Local Similarity 100.0%; Pred. No. 1.6e-42;  
Matches 85; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 209 EPVDPRLPWPVKHPSQPKTACTNCYKCCFHCQVCFITKALGISYGRKKRRRPPQG 268  
|||||  
Db 2 EPVDPRLPWPVKHPSQPKTACTNCYKCCFHCQVCFITKALGISYGRKKRRRPPQG 61

QY 269 SOTHQVSLSKOPTSOSRGDPTGPK 293  
|||||  
Db 62 SOTHQVSLSKOPTSOSRGDPTGPK 86

RESULT 18  
PCT-US93-06077-2  
Sequence 2, Application PC/TUS9306077  
GENERAL INFORMATION:  
APPLICANT: Immunobiology Research, Institute Inc.  
TITLE OF INVENTION: Vaccine Interdiction of Extracellular  
TITLE OF INVENTION: Transactivating Proteins of Human Immunodeficiency Virus  
TITLE OF INVENTION: and Other Chronically Infecting Viruses Employing Similar  
NUMBER OF SEQUENCES: 38  
CORRESPONDENCE ADDRESS:  
ADDRESS: Howson and Howson  
STREET: Spring House Corporate Cntr, PO Box 457  
CITY: Spring House  
STATE: Pennsylvania  
COUNTRY: USA  
ZIP: 19477  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: PatentIn Release #1.0, Version #1.25  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: PCT/US93/06077  
FILING DATE:  
CLASSIFICATION:  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: US 08/247.991  
FILING DATE: 23-MAY-1994  
ATTORNEY/AGENT INFORMATION:  
NAME: Bak, Mary E.  
REGISTRATION NUMBER: 31,215  
REFERENCE/DOCKET NUMBER: IRI44PCT  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: 215-540-9200

TELEFAX: 215-540-5818  
; INFORMATION FOR SEQ ID NO: 2:  
; SEQUENCE CHARACTERISTICS:  
; LENGTH: 86 amino acids  
; TYPE: amino acid  
; TOPOLOGY: linear  
; MOLECULE TYPE: protein  
PCT-US95-06077-2

Query Match 28.8%; Score 486; DB 5; Length 86;  
Best Local Similarity 100.0%; Pred. No. 1.6e-42;  
Matches 85; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 209 EPVDPRLPEPKHPSGPKACTNCKCKCFHCQVCFITKALGISYGRKKRRRRPPQG 268  
Db 2 EPVDPRLPEPKHPSGPKACTNCKCKCFHCQVCFITKALGISYGRKKRRRRPPQG 61

QY 269 SOTHQVSLSKOPTSQSGDPTGPK 293  
Db 62 SOTHQVSLSKOPTSQSGDPTGPK 86

RESULT 19  
US-08-815-809-5  
; Sequence 5, Application US/08815809  
; Patent No. 6004777  
; GENERAL INFORMATION:  
; APPLICANT: TARTAGLIA, James  
; APPLICANT: GOEBEL, Scott J.  
; APPLICANT: COX, William I.  
; APPLICANT: GETTIG, Russell R.  
; APPLICANT: PINCUS, Steven E.  
; APPLICANT: PAOLETTI, Enzo  
; APPLICANT: JACOBS, Bertram L.  
; TITLE OF INVENTION: VECTORS HAVING ENHANCED EXPRESSION, AND METHODS OF  
; FILE REFERENCE: 454310-3010  
; CURRENT APPLICATION NUMBER: US/08/815,809  
; CURRENT FILING DATE: 1997-03-12  
; NUMBER OF SEQ ID NOS: 23  
; SOFTWARE: PatentIn Ver. 2.0  
; SEQ ID NO 5  
; LENGTH: 266  
; TYPE: PRP  
; ORGANISM: Vaccinia virus  
US-08-815-809-5

Query Match 28.3%; Score 478.5; DB 3; Length 266;  
Best Local Similarity 77.3%; Pred. No. 4.5e-41;  
Matches 92; Conservative 4; Mismatches 6; Indels 17; Gaps 4;

QY 66 VGFPVTPQVPLRPMTYKAAVDLSHFLKEGGLIHSQRRQDILDWIYHTQGYFPDQW 125  
Db 136 VGFPVTPQVPLRPMTYKAAVDLSHFLKEGGLIHSQRRQDILDWIYHTQGYFPDQW 195

QY 126 NYTPGPGVRYPLTFGWCYKLVPEPKVEANKGENTSLHLPVSLH-GMDDPEREVLEW 183  
Db 196 NYTPGPGVRYPLTFGWCYKLVPM-----IETV-----PVKLPKMGDGP--KVKQW 238

RESULT 20  
US-08-816-155B-43  
; Sequence 43, Application US/08816155B  
; Patent No. 5990091  
; GENERAL INFORMATION:  
; APPLICANT: TARTAGLIA, JAMES  
; APPLICANT: COX, WILLIAM I.  
; APPLICANT: GETTIG, RUSSELL R.  
; APPLICANT: MARTINEZ, HECTOR  
; APPLICANT: PAOLETTI, ENZO  
; APPLICANT: PINCUS, STEVEN E.

; TITLE OF INVENTION: VECTORS HAVING ENHANCED EXPRESSION, AND  
; METHODS OF MAKING AND USES THEREOF

; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: FROMMER LAWRENCE & HAUG LLP  
; STREET: 745 FIFTH AVENUE  
; CITY: NEW YORK  
; STATE: NEW YORK  
; COUNTRY: USA  
; ZIP: 10151  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Floppy disk  
; COMPUTER: IBM PC compatible

; TITLE OF INVENTION: VECTORS HAVING ENHANCED EXPRESSION, AND  
; METHODS OF MAKING AND USES THEREOF

; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: FROMMER LAWRENCE & HAUG LLP  
; STREET: 745 FIFTH AVENUE  
; CITY: NEW YORK  
; STATE: NEW YORK  
; COUNTRY: USA  
; ZIP: 10151  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Floppy disk  
; COMPUTER: IBM PC compatible

Query Match 28.3%; Score 478.5; DB 2; Length 280;  
Best Local Similarity 77.3%; Pred. No. 4.8e-41;  
Matches 92; Conservative 4; Mismatches 6; Indels 17; Gaps 4;

QY 66 VGFPVTPQVPLRPMTYKAAVDLSHFLKEGGLIHSQRRQDILDWIYHTQGYFPDQW 125  
Db 150 VGFPVTPQVPLRPMTYKAAVDLSHFLKEGGLIHSQRRQDILDWIYHTQGYFPDQW 209

QY 126 NYTPGPGVRYPLTFGWCYKLVPEPKVEANKGENTSLHLPVSLH-GMDDPEREVLEW 183  
Db 210 NYTPGPGVRYPLTFGWCYKLVPM-----IETV-----PVKLPKMGDGP--KVKQW 252

RESULT 21  
US-09-079-587-43  
; Sequence 43, Application US/09079587  
; Patent No. 6130066  
; GENERAL INFORMATION:  
; APPLICANT: TARTAGLIA, JAMES  
; APPLICANT: COX, WILLIAM I.  
; APPLICANT: GETTIG, RUSSELL R.  
; APPLICANT: MARTINEZ, HECTOR  
; APPLICANT: PAOLETTI, ENZO  
; APPLICANT: PINCUS, STEVEN E.  
; TITLE OF INVENTION: VECTORS HAVING ENHANCED EXPRESSION, AND  
; METHODS OF MAKING AND USES THEREOF

; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: FROMMER LAWRENCE & HAUG LLP  
; STREET: 745 FIFTH AVENUE  
; CITY: NEW YORK  
; STATE: NEW YORK  
; COUNTRY: USA  
; ZIP: 10151  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Floppy disk  
; COMPUTER: IBM PC compatible

OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: PatentIn Release #1.0, Version #1.30  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/09/079,587  
FILING DATE:  
CLASSIFICATION:  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: US 08/816,155  
FILING DATE: 12-MAR-1997  
ATTORNEY/AGENT INFORMATION:  
NAME: KOWALSKI, THOMAS J.  
REGISTRATION NUMBER: 32,147  
REFERENCE/DOCKET NUMBER: 454310-2990  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: 212-588-0800  
TELEFAX: 212-588-0500  
INFORMATION FOR SEQ ID NO: 43:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 280 amino acids  
TYPE: amino acid  
STRANDEDNESS: n/a  
TOPOLOGY: linear  
MOLECULE TYPE: amino acid  
US-09-079-587-43

Query Match 28.3%; Score 478.5; DB 4; Length 280;  
Best Local Similarity 77.3%; Pred. No. 4.8e-41;  
Matches 92; Conservative 4; Mismatches 6; Indels 17; Gaps 4;  
QY 66 VGFPPVTPVPLRPMTYKAAVDLSHFLKKGGLHSGRRQDILDLDLWYHTQGYFPDQW 125  
DB 150 VGFPPVTPVPLRPMTYKAAVDLSHFLKKGGLHSGRRQDILDLDLWYHTQGYFPDQW 209  
QY 126 NYTPGGVRYPLTFGCYKLVPEPKVEANKGENTSLHPVSLH-GMDDPEREVLEW 183  
DB 210 NYTPGGVRYPLTFGCYKLVPEPKVEANKGENTSLHPVSLH-GMDDPEREVLEW 252

RESULT 22  
US-08-505-210-1  
Sequence 1, Application US/08505210  
Patent No. 5981258  
GENERAL INFORMATION:  
APPLICANT: MEHTALI, Majid  
APPLICANT: GUSS, Tania  
TITLE OF INVENTION: COMPOSITION OF TRANS-DOMINANT VARIANTS  
TITLE OF INVENTION: OF VIRAL PROTEINS FOR OBTAINING AN ANTIVIRAL EFFECT  
NUMBER OF SEQUENCES: 7  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: BURNS, DOANE, SWECKER & MATHIS, L.L.P.  
STREET: 1737 King Street, Suite 500  
CITY: Alexandria  
STATE: Virginia  
COUNTRY: United States  
ZIP: 22314-2756  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: PatentIn Release #1.0, Version #1.30  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/505,210  
FILING DATE: 14-AUG-1995  
CLASSIFICATION: 435  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: US 08/215,248  
FILING DATE: 21-MAR-1994  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: WO PCT/FR94/01457  
FILING DATE: 13-DEC-1994  
ATTORNEY/AGENT INFORMATION:  
NAME: Dadio, Susan M.

REGISTRATION NUMBER: 40,373  
REFERENCE/DOCKET NUMBER: 017753-066  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: (703) 836-6620  
TELEFAX: (703) 836-2021  
INFORMATION FOR SEQ ID NO: 1:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 86 amino acids  
TYPE: amino acid  
STRANDEDNESS: single  
TOPOLOGY: linear  
MOLECULE TYPE: protein  
HYPOTHETICAL: NO  
ANTI-SENSE: NO  
ORIGINAL SOURCE:  
ORGANISM: Human immunodeficiency virus type 1  
STRAIN: Lai  
INDIVIDUAL ISOLATE: sequence of the TAT protein of HIV-1  
US-08-505-210-1

Query Match 27.8%; Score 470; DB 2; Length 86;  
Best Local Similarity 96.5%; Pred. No. 7.1e-41;  
Matches 82; Conservative 0; Mismatches 3; Indels 0; Gaps 0;  
QY 209 EPVDPRLPEPKHPSQPKTACTCYCKKCCFHCQVCFITKALGISYGRKKRRRRPPOG 268  
DB 2 EPVDPRLPEPKHPSQPKTACTCYCKKCCFHCQVCFITKALGISYGRKKRRRRPPOG 61  
QY 269 SOTHQVSLSKQPTSQSRGDPGTGPKE 293  
DB 62 SOTHQVSLSKQPTSQSRGDPGTGPKE 86

RESULT 23  
US-09-099-333-1  
Sequence 1, Application US/09099333A  
Patent No. 6228369  
GENERAL INFORMATION:  
APPLICANT: MEHTALI, Majid  
APPLICANT: GUSS, Tania  
TITLE OF INVENTION: COMPOSITION OF TRANS-DOMINANT VARIANTS OF VIRAL  
TITLE OF INVENTION: PROTEINS FOR OBTAINING AN ANTI-VIRAL EFFECT  
FILE REFERENCE: 017753-091  
CURRENT APPLICATION NUMBER: US/09/099,333A  
CURRENT FILING DATE: 1998-06-18  
EARLIER APPLICATION NUMBER: FR 93 14914  
EARLIER FILING DATE: 1993-12-13  
EARLIER APPLICATION NUMBER: US 08/215,248  
EARLIER FILING DATE: 1994-03-21  
EARLIER APPLICATION NUMBER: US 08/505,210  
EARLIER FILING DATE: 1995-08-14  
NUMBER OF SEQ ID NOS: 7  
SOFTWARE: PatentIn Ver. 2.0  
SEQ ID NO 1  
LENGTH: 86  
TYPE: PRT  
ORGANISM: Human immunodeficiency virus type 1  
US-09-099-333-1

Query Match 27.8%; Score 470; DB 4; Length 86;  
Best Local Similarity 96.5%; Pred. No. 7.1e-41;  
Matches 82; Conservative 0; Mismatches 3; Indels 0; Gaps 0;  
QY 209 EPVDPRLPEPKHPSQPKTACTCYCKKCCFHCQVCFITKALGISYGRKKRRRRPPOG 268  
DB 2 EPVDPRLPEPKHPSQPKTACTCYCKKCCFHCQVCFITKALGISYGRKKRRRRPPOG 61  
QY 269 SOTHQVSLSKQPTSQSRGDPGTGPKE 293  
DB 62 SOTHQVSLSKQPTSQSRGDPGTGPKE 86

RESULT 24  
US-08-417-210A-96  
; Sequence 96, Application US/08417210A  
; Patent No. 5863542  
; GENERAL INFORMATION:  
; APPLICANT: PAOLETTI, ENZO  
; APPLICANT: TARTAGLIA, JAMES  
; APPLICANT: COX, WILLIAM I.  
; TITLE OF INVENTION: IMMUNODEFICIENCY RECOMBINANT POXVIRUS  
; NUMBER OF SEQUENCES: 148  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: CURTIS, MORRIS & SAFFORD, P.C.  
; STREET: 530 FIFTH AVENUE  
; CITY: NEW YORK  
; STATE: NEW YORK  
; COUNTRY: USA  
; ZIP: 10036  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Floppy disk  
; OPERATING SYSTEM: IBM PC compatible  
; SOFTWARE: PC-DOS/MS-DOS  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/08/417,210A  
; FILING DATE: 05-APR-1995  
; CLASSIFICATION: 435  
; ATTORNEY/AGENT INFORMATION:  
; NAME: KOWALSKI, THOMAS J.  
; REGISTRATION NUMBER: 32,147  
; REFERENCE/DOCKET NUMBER: 454310-2690  
; TELECOMMUNICATION INFORMATION:  
; TELEPHONE: 212-840-3333  
; INFORMATION FOR SEQ ID NO: 96:  
; SEQUENCE CHARACTERISTICS:  
; LENGTH: 83 amino acids  
; TYPE: amino acid  
; STRANDEDNESS: single  
; TOPOLOGY: linear  
; MOLECULE TYPE: amino acid  
; US-08-417-210A-96

Query Match 27.6%; Score 466; DB 2; Length 83;  
Best Local Similarity 100.0%; Pred. No. 1.7e-40;  
Matches 82; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 66 VGFVPTQVPLRPMYKAAVDLSHFLKEGGLGSLHSQRQDILDWYHTQGYFFPDWQ 125  
|||||  
Db 2 VGFVPTQVPLRPMYKAAVDLSHFLKEGGLGSLHSQRQDILDWYHTQGYFFPDWQ 61  
|||||

QY 126 NYTPGPGVRYPLTFGCYKLVLP 147  
|||||  
Db 62 NYTPGPGVRYPLTFGCYKLVLP 83  
|||||

RESULT 25  
US-08-053-079A-15  
; Sequence 15, Application US/08053079A  
; Patent No. 5606026  
; GENERAL INFORMATION:  
; APPLICANT: Rodman  
; TITLE OF INVENTION: Natural Human IgM Antibodies  
; NUMBER OF SEQUENCES: 15  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: Darby & Darby  
; STREET: 805 Third Ave.  
; CITY: New York  
; STATE: New York  
; COUNTRY: U.S.A.  
; ZIP: 10022  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Floppy disk

; COMPUTER: IBM PC compatible  
; OPERATING SYSTEM: PC-DOS/MS-DOS  
; SOFTWARE: Wordperfect 5.1  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/08/053,079A  
; FILING DATE: 26-APR-1993  
; CLASSIFICATION: 530  
; ATTORNEY/AGENT INFORMATION:  
; NAME: ROBINSON, JOSEPH, R.  
; REGISTRATION NUMBER: 33,448  
; REFERENCE/DOCKET NUMBER: 4436/16060US4  
; TELECOMMUNICATION INFORMATION:  
; TELEPHONE: (212)527-7700  
; TELEFAX: (212)753-6237  
; TELEX: 236687  
; INFORMATION FOR SEQ ID NO: 15:  
; SEQUENCE CHARACTERISTICS:  
; LENGTH: 82  
; TYPE: amino acid  
; STRANDEDNESS: single  
; TOPOLOGY: linear  
; US-08-053-079A-15

Query Match 27.0%; Score 455; DB 1; Length 82;  
Best Local Similarity 98.8%; Pred. No. 2.3e-39;  
Matches 80; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 209 EPVDPRLPEPKHSGSQPTACTNCYCKKCFHCQVCFITKALGISYGRKKRRRRPQ 268  
|||||  
Db 2 EPVDPRLPEPKHSGSQPTACTNCYCKKCFHCQVCFITKALGISYGRKKRRRRPQ 61  
|||||

QY 269 SQTHQVSLSKQPTSQSRGDPT 289  
|||||  
Db 62 SQTHQVSLSKQPTSQSRGDPT 82  
|||||

RESULT 26  
US-08-659-251-4  
; Sequence 4, Application US/08659251  
; Patent No. 5883081  
; GENERAL INFORMATION:  
; APPLICANT: Kraus, Guenter  
; APPLICANT: Wong-Staal, Flossie  
; APPLICANT: Talbott, Randy  
; APPLICANT: Posschla, Eric  
; TITLE OF INVENTION: Isolation of No. 5883081el HIV-2 Proviruses  
; NUMBER OF SEQUENCES: 50  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: Townsend and Townsend and Crew LLP  
; STREET: Two Embarcadero Center, Eighth Floor  
; CITY: San Francisco  
; STATE: California  
; COUNTRY: USA  
; ZIP: 94111-3834  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Floppy disk  
; OPERATING SYSTEM: IBM PC compatible  
; SOFTWARE: PC-DOS/MS-DOS  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/08/659,251  
; FILING DATE: No. 5883081 yet assigned  
; CLASSIFICATION: 514  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: US 60/001,441  
; FILING DATE: 26-JUL-1995  
; ATTORNEY/AGENT INFORMATION:  
; NAME: Garrett-Wackowski, Eugenia  
; REGISTRATION NUMBER: 37,330  
; REFERENCE/DOCKET NUMBER: 02307E-056410US  
; TELECOMMUNICATION INFORMATION:  
; TELEPHONE: (415) 576-0200



INFORMATION FOR SEQ ID NO: 4:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 253 amino acids  
TYPE: amino acid  
STRANDEDNESS:  
TOPOLOGY: linear  
MOLECULE TYPE: protein  
FEATURE:  
NAME/KEY: Protein  
LOCATION: 1..253  
OTHER INFORMATION: /note= "nef protein encoded by HIV-2KR"  
PCT-US96-11445-4

Query Match 25.5%; Score 430.5; DB 5; Length 253;  
Best Local Similarity 39.4%; Pred. No. 3.4e-36;  
Matches 99; Conservative 29; Mismatches 82; Indels 41; Gaps 6;  
QY 1 MGKSKSVGWGTPVRRMRRAEPAADG---VGAARDLE-KHGAITSSNTAATNA-- 54  
DB 1 MGASGKCSRLQGLRERLLRARGTCGGQWDGSGAGEYLQFQSGRGONLPSCGQRY 60  
QY 55 -----CAW-----LEAQEEVEGFPVTPQVPLRPMTYKAAVDLS 120  
DB 61 QQGFDMNTWRTPAAGRECTLYKQONMDVDADNDNLIGVPTPRVPLRAMTYKLAVDIS 120  
QY 89 HFLKEGGLGLHSORRODILDLWYHTQGYFPDQWNTPGGVRYPLTFGWCYKLVVPV 148  
DB 121 HFLNKGGLDGYSERRRIIDIYMEKEGIPDQWNTHTGPGVRYKFFGWLKLVVPV 180  
QY 149 EPDKVEEANKGENTSLHPSVLSHGMDPERVLEWRFDLSRLAFHHVARELHPEYFNKCTS 208  
DB 181 DVPQGE-----EDCHLLHPAQTSGGDDPHGELTLMRFPRLAYEYAFNRYPEFGYKSG 235  
QY 209 EPVDPRLPEPK 219  
DB 236 LPE-----EEWK 242

RESULT 29  
US-09-030-613-17  
Sequence 17, Application US/09030613  
Patent No. 6083706  
GENERAL INFORMATION:  
APPLICANT: Florckiewicz, Robert Z.  
APPLICANT: Baird, J. Andrew  
TITLE OF INVENTION: INHIBITORS OF LEADERLESS PROTEIN EXPORT  
NUMBER OF SEQUENCES: 36  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: SEED AND BERRY LLP  
STREET: 6300 Columbia Center, 701 Fifth Avenue  
CITY: Seattle  
STATE: Washington  
COUNTRY: USA  
ZIP: 98104  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: Patent In Release #1.0, Version #1.30  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/09/030,613  
FILING DATE: 25-FEB-1998  
CLASSIFICATION:  
ATTORNEY/AGENT INFORMATION:  
NAME: No. 6083706tenburg Ph.D., Carol  
REGISTRATION NUMBER: 39,317  
REFERENCE/DOCKET NUMBER: 760100.418C1  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: (206) 622-4900  
TELEFAX: (206) 682-6031  
INFORMATION FOR SEQ ID NO: 17:  
SEQUENCE CHARACTERISTICS:

LENGTH: 72 amino acids  
TYPE: amino acid  
STRANDEDNESS:  
TOPOLOGY: linear  
US-09-030-613-17  
Query Match 24.2%; Score 409; DB 3; Length 72;  
Best Local Similarity 100.0%; Pred. No. 9.7e-35;  
Matches 71; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
QY 209 EPVDPRLPEPKHPSGQPKTACTNCTCYKCCFHCQVCFITKALGISYGRKKRRRRPPQG 268  
DB 2 EPVDPRLPEPKHPSGQPKTACTNCTCYKCCFHCQVCFITKALGISYGRKKRRRRPPQG 61  
QY 269 SOTHQVSLSKQ 279  
DB 62 SOTHQVSLSKQ 72  
RESULT 30  
US-09-451-905-17  
Sequence 17, Application US/09451905  
Patent No. 6306613  
GENERAL INFORMATION:  
APPLICANT: Robert Z. Florckiewicz  
APPLICANT: Andrew Baird  
APPLICANT: Dale E. Warnock  
TITLE OF INVENTION: MODULATORS OF LEADERLESS PROTEIN EXPORT  
FILE REFERENCE: 200124.402C4  
CURRENT APPLICATION NUMBER: US/09/451,905  
CURRENT FILING DATE: 1999-12-01  
NUMBER OF SEQ ID NOS: 48  
SOFTWARE: FastSeq for Windows Version 4.0  
SEQ ID NO 17  
LENGTH: 72  
TYPE: PRT  
ORGANISM: Homo sapien  
US-09-451-905-17  
Query Match 24.2%; Score 409; DB 4; Length 72;  
Best Local Similarity 100.0%; Pred. No. 9.7e-35;  
Matches 71; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
QY 209 EPVDPRLPEPKHPSGQPKTACTNCTCYKCCFHCQVCFITKALGISYGRKKRRRRPPQG 268  
DB 2 EPVDPRLPEPKHPSGQPKTACTNCTCYKCCFHCQVCFITKALGISYGRKKRRRRPPQG 61  
QY 269 SOTHQVSLSKQ 279  
DB 62 SOTHQVSLSKQ 72

Search completed: August 26, 2002, 08:10:39  
Job time: 128 sec